Statistical Analysis Plan Cover Page

ACURATE IDE: Transcatheter Replacement of Stenotic Aortic Valve through Implantation of ACURATE in Subjects InDicatEd for TAVR

ACURATE IDE

S2408

National Clinical Trial Identification Number:
NCT03735667

Document Date
September 15, 2023

Statistical Analysis Plan

ACURATE IDE: Transcatheter Replacement of Stenotic Aortic Valve through Implantation of ACURATE in Subjects InDicatEd for TAVR

ACURATE IDE

S2408

APPROVALS (Check/Complete one below):

☒ Approvals are captured electronically

TABLE OF CONTENTS

1 PROTOCOL SUMMARY

<table>
<tr><td>Study Objective(s)</td><td>To evaluate safety and effectiveness of the ACURATE Transfemoral Aortic Valve System for transcatheter aortic valve replacement (TAVR) in subjects with severe native aortic stenosis who are indicated for TAVR.</td></tr>
<tr><td>Planned Indication(s) for Use</td><td>The ACURATE Transfemoral Aortic Valve System is intended to improve aortic valve function in subjects with severe native aortic stenosis who are indicated for TAVR.</td></tr>
<tr><td>Test Device and Sizes</td><td>The ACURATE Transfemoral Aortic Valve System (ACURATE) consists of the following.
<table>
<tr><th>Device Name/Size</th><th>Description</th></tr>
<tr><td>ACURATE neo2™ Aortic Valve
Valve size:
- S (small)
- M (medium)
- L (large)
with 23mm, 25mm, and 27mm nominal diameter at waist level, respectively</td><td>Includes 3 main components:
- A three-leaflet porcine pericardial bioprosthetic aortic valve;
- A self-expandable Nitinol stent;
- A double porcine pericardium skirt sutured on the inner and outer surface of the stent to prevent paravalvular leaks.
Introduced via the iliofemoral artery.</td></tr>
<tr><td>ACURATE neo2™ Transfemoral Delivery System
The delivery system is compatible with the S, M, and L valve sizes.</td><td>Allows positioning and delivery of the transcatheter valve via iliofemoral access.</td></tr>
<tr><td>ACURATE Prime™ Aortic Valve XL*
Valve size:
- XL (extra-large)
with 29mm nominal diameter at waist level</td><td>Includes 3 main components:
- A three-leaflet porcine pericardial bioprosthetic aortic valve.
- A self-expandable Nitinol stent.
- A double porcine pericardium skirt sutured on the inner and outer surface of the stent to prevent paravalvular leaks.
Introduced via the iliofemoral artery.</td></tr>
<tr><td>ACURATE Prime™ Transfemoral Delivery System XL*</td><td>Allows positioning and delivery of the transcatheter XL valve via iliofemoral access.</td></tr>
<tr><td>ACURATE Prime™ Loading Kit XL*</td><td>Allows loading of the XL valve onto the delivery system.</td></tr>
<tr><td colspan="2">* ACURATE Prime XL is only available in the United States.</td></tr>
</table>
</td></tr>
</table>

Control Device and Sizes	In the Main Randomized Cohort and in the Extended Durability Study (see study cohorts described below), subjects assigned to the control arm will receive a commercially available balloon-expandable SAPIEN 3™ Transcatheter Heart Valve or future iteration (SAPIEN 3; Edwards Lifesciences LLC, Irvine, CA, USA) or a commercially available self-expanding CoreValve® Transcatheter Aortic Valve Replacement System, CoreValve® Evolut™ R Recapturable TAVR System, EVOLUT™ PRO System, or future iteration (CoreValve; Medtronic, Inc., Dublin, Ireland) TAVR device that is introduced via the femoral artery using conventional catheterization techniques. **Note 1:** Every subject in the Main Randomized Cohort and the Extended Durability Study must be deemed treatable with an available size of both the test (ACURATE) device and the control device.
Study Design	ACURATE IDE is a prospective, multicenter trial designed to evaluate the safety and effectiveness of the ACURATE Transfemoral Aortic Valve System for TAVR in subjects who have severe native aortic stenosis and are indicated for TAVR. Study cohorts include the following: • **Main Randomized Cohort:** A prospective, multicenter, 1:1 randomized controlled trial (RCT; ACURATE versus Control [commercially available SAPIEN 3 or CoreValve] TAVR device). Randomization will be stratified by center and by intended control device. • **Roll-In Cohort:** A non-randomized roll-in phase with the test device. Centers that do not have implantation experience with the ACURATE *neo*™ Aortic Bioprosthesis (transfemoral delivery; Boston Scientific Corporation, Marlborough, MA, USA) will perform at least 2 roll-in cases with ACURATE *neo2* before commencing treatment in the randomized cohort. Centers with prior experience with ACURATE are not required to do roll-in cases. Data from roll-in subjects will be summarized separately from the randomized cohort and will not be included in the primary endpoint analysis. • **4D CT Imaging Substudy:** Selected centers with the ability to perform high quality 4D computed tomography (CT) scans will include subjects from the Main Randomized Cohort in a 4D CT Imaging Substudy to assess the prevalence of reduced leaflet mobility and hypoattenuated leaflet thickening (HALT) and the relationship, if any, to clinical events. Subjects will be randomized to test (ACURATE) and control device.

- **ACURATE *Prime*™ XL Nested Registry:** A non-randomized, nested registry cohort of subjects who will receive the ACURATE *Prime*™ Transfemoral Aortic Valve System XL (ACURATE *Prime* XL Nested Registry). Participating centers will be a subset of United States centers that have enrolled subjects in ACURATE IDE. Data from subjects in this nested registry will be summarized separately from other cohorts.
- **ACURATE Extended Durability Study:** An additional 1:1 randomized study (ACURATE versus Control [commercially available SAPIEN 3 or CoreValve] TAVR device) including only subjects considered to be at low surgical risk. Subjects will receive ACURATE *neo2* (S, M, or L valve sizes) or ACURATE *Prime* XL. Randomization will be stratified by center and by intended control device. Low-risk subjects receiving ACURATE *neo2* will be enrolled in the Extended Durability Study only after enrollment of the Main Randomized Cohort is completed. Enrollment of low risk subjects with ACURATE *Prime* XL will start after enrollment in both the Main Randomized Cohort and the ACURATE *Prime* XL Nested Registry is completed. Data from subjects in the Extended Durability Study will be summarized separately from other cohorts.
  Note: Centers must complete the roll-in phase of the study, if applicable, before participating in the ACURATE Extended Durability Study.
- **ACURATE Continued Access Study (CAS):** An additional cohort of subjects receiving ACURATE *neo2* (S, M, and L valve sizes) or ACURATE *Prime* XL. Enrollment of subjects with ACURATE *neo2* will start after enrollment of the ACURATE IDE Main Randomized Cohort is completed. Enrollment of subjects with ACURATE *Prime* XL will start after enrollment in both the Main Randomized Cohort and the ACURATE *Prime* XL Nested Registry is completed. Subjects at all surgical risks may be enrolled in ACURATE CAS, but low-risk subjects will be considered for enrollment only after enrollment in the Extended Durability Study is completed. Data from subjects in the ACURATE CAS will be summarized separately from other cohorts and will be used to further assess performance and safety.
  Note: Centers must complete the roll-in phase of the study, if applicable, before participating in ACURATE CAS.

The devices and risk levels for the ACURATE IDE cohorts are summarized in the figure below.

Devices and Risk Levels for ACURATE IDE Cohorts

Device	Cohort				
	Roll-In	**Main RCT**	***Prime* XL Nested Registry**	**Extended Durability [1]**	**Continued Access Study[1]**
ACURATE *neo2* (S, M, L)	All risks	All risks	N/A	Low risk	All risks[2]
ACURATE *Prime* (XL)	N/A	N/A	All risks	Low risk[3]	All risks[2,3]

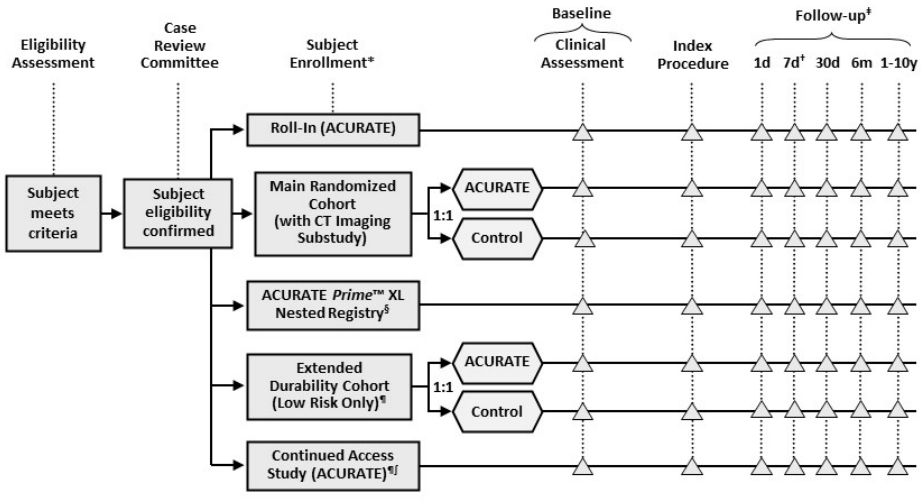
Note 1: Enrollment of subjects in the Extended Durability cohort and the CAS cohort will begin after completion of enrollment in the Main RCT.

Note 2: Enrollment of low-surgical-risk subjects in the CAS cohort will begin after completion of enrollment the Extended Durability cohort.

Note 3: Enrollment of subjects receiving ACURATE *Prime* XL in the Extended Durability and CAS cohorts will begin after enrollment completion in both the Main RCT and *Prime* XL Nested Registry.

Abbreviations: CAS=Continued Access Study; N/A=not applicable; RCT=randomized controlled trial

The ACURATE IDE study will be conducted in accordance with 21 CFR Parts 11, 50, and 54; the relevant parts of the International Conference on Harmonisation (ICH) Guidelines for Good Clinical Practices (GCP); the International Standard ISO 14155 Clinical Investigation of Medical Devices for Human Subjects – Good Clinical Practice; ethical principles that have their origins in the Declaration of Helsinki; and pertinent individual country/state/local laws and regulations. The study shall not begin until the required approval/favorable opinion from the Institutional Review Board/Research Ethics Board/Human Research Ethics Committee/ Independent Ethics Committee (IRB/REB/HREC/IEC) and/or regulatory authority has been obtained, if appropriate. The study design is summarized in the figure below.



* Randomized subjects are considered enrolled upon randomization; roll-in, ACURATE *Prime* XL Nested Registry, and Continued Access Study subjects are considered enrolled when an attempt is made to insert the ACURATE Transfemoral Aortic Valve System into the femoral artery.

† Discharge or 7 days, whichever comes first

‡ Includes clinic/in-person visit or telehealth visit (https://telehealth.HHS.gov) at 30 days, 6 months, and years 1 – 5, 7 & 10; telephone follow-up at years 6, 8 & 9. For subjects in the CT Imaging Substudy, the visits at 30 days and 1 year must be done in the clinic (or in-hospital). Enrolled subjects who do not have a study device implanted will be assessed through 1-year post procedure for safety.

§ Includes subjects who receive the ACURATE *Prime*™ XL device.

¶ Includes subjects who receive the ACURATE *neo2*™ (S, M, L) device (after conclusion of enrollment in the Main Randomized Cohort) or ACURATE *Prime*™ XL device (after conclusion of XL Nested Registry enrollment).

∫ Low-risk subjects will be considered for enrollment in the Continued Access Study after enrollment is completed in the Extended Durability Study.

ACURATE IDE Study Design Overview

Planned Number of Subjects and Planned Number of Investigational Centers/ Countries

Subjects will be enrolled at up to 85 centers in the United States, Canada, Europe, and Australia. In total, there will be up to 2820 subjects enrolled in ACURATE IDE across the various cohorts. It is expected that >80% of enrolled subjects will be from the United States. The number of subjects intended for each cohort is shown in the table below.

To achieve sufficient distribution of lower risk subjects in the Main Randomized Cohort, there will be ≥ 30% intermediate risk randomized subjects and ≥ 35% low risk randomized subjects enrolled. The ACURATE *Prime* XL Nested Registry and the ACURATE Continued Access Study will enroll subjects at all surgical risks. Only low-risk subjects will be enrolled in the ACURATE Extended Durability Cohort.

Cohort	Number of Subjects
Main Randomized Cohort	1,500*
Roll-In Cohort	Up to 170
ACURATE *Prime* XL Nested Registry	50‡
ACURATE Extended Durability Study	A minimum of 100†
ACURATE Continued Access Study (CAS)	Up to 1000§

***** A subset of subjects (minimum of 200) in the Main Randomized Cohort will also be enrolled in the 4D CT Imaging Substudy.

‡ Subjects will be enrolled at up to 20 centers in the United States.

	† Only low-risk subjects will be included in the Extended Durability Study. § Low-risk subjects will be considered for enrollment in the CAS after enrollment is completed in the Extended Durability Study
Primary Endpoint	Composite of all-cause mortality, all stroke, and rehospitalization† at 1 year. † Hospitalization for valve-related symptoms or worsening congestive heart failure (NYHA class III or IV); per VARC-2 definition
Additional Measurements	This section describes required assessments to be performed in the ACURATE IDE study. Additional measurements based on the VARC[a] endpoints and definitions will be collected peri- and post-procedure, at discharge or 7 days post-procedure (whichever comes first), 30 days, 6 months, and annually through 10 years post index procedure, unless otherwise specified below. Safety endpoints adjudicated by an independent Clinical Events Committee (CEC): • Mortality: all-cause, cardiovascular, and non-cardiovascular • Stroke: disabling and non-disabling • Myocardial infarction (MI): periprocedural (≤72 hours post index procedure) and spontaneous (>72 hours post index procedure) • Bleeding: life-threatening (or disabling) and major (through 5 years) • Acute kidney injury (≤7 days post index procedure): based on the AKIN System Stage 3 (including renal replacement therapy) or Stage 2 • Major vascular complication (through 5 years) • Repeat procedure for valve-related dysfunction (surgical or interventional therapy) • Hospitalization for valve-related symptoms or worsening congestive heart failure (NYHA class III or IV) • New permanent pacemaker implantation resulting from new or worsened conduction disturbances • New onset of atrial fibrillation or atrial flutter • Coronary obstruction: periprocedural (≤72 hours post index procedure) • Ventricular septal perforation: periprocedural (≤72 hours post index procedure) • Mitral apparatus damage: periprocedural (≤72 hours post index procedure)

	• Cardiac tamponade: periprocedural (≤72 hours post index procedure) • Valve migration • Valve embolization • Ectopic valve deployment • Transcatheter aortic valve (TAV)-in-TAV deployment • Prosthetic aortic valve thrombosis • Prosthetic aortic valve endocarditis • Device Performance endpoints peri- and post-procedure: - Successful vascular access, delivery and deployment of the study valve, and successful retrieval of the delivery system - Grade of aortic valve regurgitation: paravalvular, central, and combined • Device success, assessed at procedure and defined as absence of procedural mortality, correct positioning of a single valve into the proper anatomical location, and intended performance of the study device (indexed effective orifice area [iEOA] >0.85 cm^2/m^2 for BMI <30 kg/cm^2 and iEOA >0.70 cm^2/m^2 for BMI ≥30 kg/cm^2 plus either a mean aortic valve gradient <20 mmHg or a peak velocity <3m/sec, and no moderate or severe prosthetic valve aortic regurgitation) • Additional indications of prosthetic aortic valve performance as measured by transthoracic echocardiography (TTE; see **Note 2** below) and assessed by an independent core laboratory, including EOA, mean and peak aortic gradients, peak aortic velocity, and grade of aortic regurgitation • Functional status as evaluated by New York Heart Association (NYHA) classification (see **Note 3** below) • Neurological status (see **Note 4** below) as determined by the following: - National Institutes of Health Stroke Scale (NIHSS) conducted by a neurology professional or certified personnel at discharge and 1 year - Modified Rankin Scale (mRS) conducted by a neurology professional or certified personnel at discharge and all follow-up visits (see **Note 4** below) - Neurological physical exam conducted by a neurologist, neurology fellow, neurology physician assistant, or neurology nurse practitioner in all subjects where stroke is suspected • Health status as evaluated by Kansas City Cardiomyopathy and SF-12 Quality of Life questionnaires at baseline, 1 month, and 1 and 5 years.

	• For subjects in the CT Imaging Substudy, assessments using 4D CT at 30 days and 1 year will be done as listed below. Data will be evaluated by an independent CT core laboratory. - Assessment of leaflet mobility - Assessment of hypoattenuated leaflet thickening (HALT) - Assessment of leaflet thrombosis **Note 2:** At least 1 echocardiogram must be obtained before discharge or 7 days (whichever comes first); if multiple echocardiographic studies are performed prior to discharge and within 7 days of the procedure, the latest study performed will be used for analysis. **Note 3:** Echocardiography and NYHA assessment are not required in years 6, 8, and 9 (telephone follow-up only). **Note 4:** The mRS is required at all follow-up visits up to 5 years. For subjects diagnosed with a stroke, a neurological physical exam, mRS, and NIHSS must be performed after the event; mRS must also be administered at 90±14 days following a stroke; the simplified mRS questionnaire may be used for this follow-up assessment. a: Leon M, *et al. J Am Coll Cardiol.* 2011;57:253–69 Kappetein AP, *et al*. *J Am Coll Cardiol*. 2012;60:1438–54 Alu M, et al. An introduction to "VARC-3: a focused update." Presented at *SHDS*, Chicago 2018.
Method of Assigning Subjects to Treatment	**Main Randomized Cohort:** A computer generated list of random treatment allocations (randomization schedule) will be used to assign subjects to treatment in a 1:1 ratio of ACURATE to Control. A subset of subjects in the randomized cohort will also be enrolled in the 4D CT Imaging Substudy. **Roll-In Cohort:** For centers that do not have implantation experience with the ACURATE *neo*™ Aortic Bioprosthesis (transfemoral delivery) at least 2 roll-in subjects will be treated before commencing treatment in the randomized cohort. Centers with prior experience with ACURATE are not required to do roll-in cases. For the main randomized cohort and the roll-in cohort, subjects will have a documented aortic annulus size of ≥21mm and ≤27mm based on pre-procedure diagnostic imaging. **ACURATE *Prime* XL Nested Registry:** Subjects will have a documented aortic annulus size of ≥26.5 mm and ≤29 mm based on pre-procedure diagnostic imaging. **ACURATE Extended Durability Study:** A computer generated list of random treatment allocations (randomization schedule) will be used to

	assign subjects to treatment in a 1:1 ratio of ACURATE to Control. Only subjects considered to be at low surgical risk will be included in the Extended Durability Study. Randomization will be stratified by center and by intended control device. Subjects will receive ACURATE *neo2* (S, M, or L valve sizes) or ACURATE *Prime* XL. Subjects with a documented aortic annulus size of ≥20.5mm and ≤27 mm may be enrolled in the Extended Durability Study after enrollment is concluded in the Main Randomized Cohort. Subjects with a documented aortic annulus size of ≥26.5mm and ≤29mm may be enrolled in the Extended Durability Study after enrollment is concluded in both the Main Randomized Cohort and the ACURATE Prime XL Nested Registry. **Note:** Centers must complete the roll-in phase of the study, if applicable, before participating in the ACURATE Extended Durability Study. **ACURATE Continued Access Study (CAS):** Subjects will receive ACURATE *neo2* (S, M, or L valve sizes) or ACURATE *Prime* XL. Subjects with a documented aortic annulus size of ≥20.5mm and ≤27mm may be enrolled in ACURATE CAS after enrollment is concluded in the Main Randomized Cohort. Subjects with a documented aortic annulus size of ≥26.5mm and ≤29mm may be enrolled in ACURATE CAS after enrollment is concluded in both the Main Randomized Cohort and the ACURATE *Prime* XL Nested Registry. Subjects at all surgical risk levels may be enrolled in ACURATE CAS but low-risk subjects will be considered for enrollment only after enrollment is completed in the Extended Durability Study. **Note:** Centers must complete the roll-in phase of the study, if applicable, before participating in ACURATE CAS.
Follow-up Schedule	All subjects implanted with a test or control device will be assessed at baseline, peri- and post-procedure, at discharge or 7 days post-procedure (whichever comes first), 30 days, 6 months, 1 year, and then annually for up to 10 years post-procedure. Subjects who are enrolled but not implanted with a test or control device at the time of the procedure will be followed for safety through 1 year. The visits at 30 days, 6 months, 1−5 years, 7 years, and 10 years are to be an office/clinical/in-person or telehealth (https/telehealth.HHS.gov) visit (see **Note 5** below) but may be done in-hospital should the subject be admitted at the time. Telephone follow-up is allowed at 6, 8, and 9 years. Procedures at each scheduled visit are described above under "Additional Measurements." **Note 5:** For subjects in the CT Imaging Substudy, the visits at 30 days and 1 year must be done in the clinic (or in-hospital).

Study Duration	Subjects implanted with a test or control device will be followed for 10 years after the index procedure. Enrollment is expected to be completed in approximately 62 months; therefore, the total study duration is estimated to be approximately 16 years.
Participant Duration	The study duration for each subject is expected to be approximately 10 years.
Adjunctive Pharmacologic Therapy	**Anticoagulant Therapy** Anticoagulant therapy (e.g., unfractionated heparin) per local standard of care must be administered during the implant procedure, with a recommended target activated clotting time of ≥250 seconds during the index procedure. **Anti-Platelet Therapy** Per US society guidelines[b], antiplatelet therapy is recommended after TAVR to decrease the risk of thrombotic or thromboembolic complications if there are no contraindications to these medications. ACURATE IDE study subjects must receive antiplatelet therapy (aspirin and/or a $P2Y_{12}$ inhibitor) for at least 1 month following valve implant (see below for recommended doses). It should be noted, however, that recent clinical evidence points to increased bleeding risk post TAVR among subjects receiving dual antiplatelet therapy or antiplatelet therapy plus anti-coagulation (among subjects indicated for anti-coagulation)[c]. Aspirin A loading dose of aspirin (recommended dose of 75–325 mg) is recommended for subjects who have not been on aspirin therapy for ≥72 hours at the time of the index procedure. The loading dose should be administered prior to the implant procedure. Subjects who have been taking aspirin daily for ≥72 hours at the time of the index procedure do not require a loading dose. After the valve implant procedure, the recommended aspirin dose is ≥75 mg daily. It is recommended that daily aspirin be given indefinitely as per local standard of care. $P2Y_{12}$ Inhibitor A loading dose of a $P2Y_{12}$ inhibitor (recommended dose of ≥300 mg for clopidogrel, 60 mg for prasugrel, 180 mg for ticagrelor) is recommended for subjects who have not been on $P2Y_{12}$ therapy for ≥72 hours at the time of the index procedure. The loading dose should be administered prior to the implant procedure.

	After the valve implant procedure, a $P2Y_{12}$ inhibitor and/or aspirin is required for at least 1 month. **Note 6:** Anti-platelet therapy dosing and the type of $P2Y_{12}$ inhibitor used should be according to the local standard of care. **Note 7:** If a subject requires chronic anticoagulation, either a $P2Y_{12}$ inhibitor or aspirin is recommended prior to and after the implant procedure in addition to the anticoagulant therapy (but both aspirin and a $P2Y_{12}$ inhibitor are not recommended). After the implant procedure, the subject should be treated with an oral anticoagulant and either a $P2Y_{12}$ inhibitor or aspirin for at least 1 month. **Note 8:** Subjects who are expected to undergo chronic anticoagulation therapy after the TAVR procedure are not eligible to be included in the 4D CT Imaging Substudy (see Additional Exclusion Criteria below). b: Otto CM, et al. *J Am Coll Cardiol.* 2021;77:e25-e197 c: Nijenhuis VJ, et al. N Engl J Med 2020;382:1696–1707 Brouwer J, et al. N Engl J Med 2020;383:1447–1457
Inclusion Criteria	IC1. Subject has documented severe symptomatic native aortic stenosis defined as follows: aortic valve area (AVA) ≤1.0 cm^2 (or AVA index ≤0.6 cm^2/m^2) AND a mean pressure gradient ≥40 mmHg, OR maximal aortic valve velocity ≥4.0 m/s, OR Doppler velocity index ≤0.25 as measured by echocardiography and/or invasive hemodynamics. **Note 9:** In cases of low flow, low gradient aortic stenosis with left ventricular dysfunction (ejection fraction <50%), dobutamine can be used to assess the grade of aortic stenosis (maximum dobutamine dose of 20 mcg/kg/min recommended)[b]; the subject may be enrolled if echocardiographic criteria are met with this augmentation. IC2. Subject has a documented aortic annulus size of ≥20.5 mm and ≤29 mm based on the center's assessment of pre-procedure diagnostic imaging (and confirmed by the Case Review Committee [CRC]) and, for the Main Randomized Cohort and the Extended Durability Study, is deemed treatable with an available size of both test and control device. IC3. For subjects with symptomatic aortic valve stenosis per IC1 definition above, functional status is NYHA Functional Class ≥ II. IC4. Heart team (which must include an experienced cardiac interventionalist and an experienced cardiac surgeon) agrees that the subject is indicated for TAVR, is likely to benefit from valve replacement, and TAVR is appropriate.

	IC5. Subject (or legal representative) understands the study requirements and the treatment procedures and provides written informed consent. IC6. Subject, family member, and/or legal representative agree(s) and subject is capable of returning to the study hospital for all required scheduled follow up visits. IC7. Subject is expected to be able to take the protocol-required adjunctive pharmacologic therapy. b: Otto CM, et al. *J Am Coll Cardiol.* 2021;77:e25-e197
Exclusion Criteria	Vulnerable subjects (ISO 14155) will not be enrolled in ACURATE IDE. EC1. Subject has a unicuspid or bicuspid aortic valve. EC2. Subject has had an acute myocardial infarction within 30 days prior to the index procedure (defined as Q-wave MI or non–Q-wave MI with total CK elevation ≥ twice normal in the presence of CK-MB elevation and/or troponin elevation). EC3. Subject has had a cerebrovascular accident or transient ischemic attack clinically confirmed by a neurologist or neuroimaging within the past 6 months prior to study enrollment. EC4. Subject is on renal replacement therapy or has eGFR <20. EC5. Subject has a pre-existing prosthetic aortic or mitral valve. EC6. Subject has severe (4+) aortic, tricuspid, or mitral regurgitation. EC7. Subject has moderate or severe mitral stenosis (mitral valve area ⩽1.5 cm^2 and diastolic pressure half-time ≥150 ms, Stage C or D[b]). EC8. Subject has a need for emergency surgery for any reason. EC9. Subject has a history of endocarditis within 6 months of index procedure or evidence of an active systemic infection or sepsis. EC10. Subject has echocardiographic evidence of new intra-cardiac vegetation or intraventricular or paravalvular thrombus requiring intervention. EC11. Subject has platelet count <50,000 cells/mm^3 or >700,000 cells/mm^3, or white blood cell count <1,000 cells/mm^3. EC12. Subject has had a gastrointestinal bleed requiring hospitalization or transfusion within the past 3 months or has other clinically significant bleeding diathesis or coagulopathy that would

	preclude treatment with required antiplatelet regimen or will refuse transfusions. EC13. Subject has known hypersensitivity to contrast agents that cannot be adequately pre-medicated or has known hypersensitivity to the protocol required medications (aspirin, all $P2Y_{12}$ inhibitors, heparin), or to the individual components of the test or control valve (nickel, titanium, stainless steel, platinum, iridium or polyethylene terephthalate [PET]). EC14. Subject has a life expectancy of less than 12 months due to non-cardiac, comorbid conditions based on the assessment of the investigator at the time of enrollment. EC15. Subject has hypertrophic cardiomyopathy. EC16. Subject has any therapeutic invasive cardiac or vascular procedure within 30 days prior to the index procedure (except for balloon aortic valvuloplasty, pacemaker implantation, or implantable cardioverter defibrillator implantation, which are allowed). EC17. Subject has untreated coronary artery disease, which in the opinion of the treating physician is clinically significant and requires revascularization. EC18. Subject has severe left ventricular dysfunction with ejection fraction <20%. EC19. Subject is in cardiogenic shock or has hemodynamic instability requiring inotropic support or mechanical support devices. EC20. Subject has arterial access that is not acceptable for the study device (test or control) delivery systems as defined in the device (test or control) Directions For Use. EC21. Subject has either of the following: • Severe vascular disease that would preclude safe access (e.g., aneurysm with thrombus that cannot be crossed safely; marked tortuosity; significant narrowing of the abdominal aorta; severe unfolding of the thoracic aorta; or thick, protruding, ulcerated atheroma in the aortic arch), OR • Severe/eccentric calcification of the aortic annulus that would prevent safe implantation of the TAVR prosthesis. EC22. Subject has current problems with substance abuse (e.g., alcohol, etc.) that may interfere with the subject's participation in this study.

	EC23. Subject is participating in another investigational drug or device study that has not reached its primary endpoint or subject intends to participate in another investigational device clinical trial within 12 months after index procedure. EC24. Subject has untreated conduction system disorder (e.g., Type II second degree atrioventricular block) that in the opinion of the treating physician is clinically significant and requires a pacemaker implantation. Enrollment is permissible after permanent pacemaker implantation. EC25. Subject has severe incapacitating dementia. b: Otto CM, et al. *J Am Coll Cardiol.* 2021;77:e25-e197
Additional Exclusion Criteria	Additional exclusion criteria apply to subjects considered for enrollment in the CT Imaging Substudy as listed below. AEC1. Subject has eGFR <30 mL/min (chronic kidney disease stage IV or stage V). AEC2. Subject has atrial fibrillation that cannot be rate controlled to ventricular response rate < 60 bpm. AEC3. Subject is expected to undergo chronic anticoagulation therapy after the index procedure. **Note 10:** Subjects treated with short-term anticoagulation post procedure can be included in the CT Imaging Substudy; in these subjects the 30-day imaging will be performed 30 days after discontinuation of anticoagulation.
Statistical Methods	
Analysis Sets	Analysis sets for the ACURATE IDE Main Randomized Cohort and the Extended Durability Study are listed below. Subjects in the Main Randomized Cohort and the Extended Durability Study are considered enrolled in the study upon randomization. - Intention-To-Treat (ITT): This population includes all subjects who sign an Informed Consent Form, are enrolled in the trial, and are randomized, whether or not an assigned study device is implanted. - Implanted: This population includes all subjects who sign an Informed Consent Form, are enrolled in the trial, and are implanted with the assigned (control versus test), randomized study device. **Note 11:** For the Main Randomized cohort and the Extended Durability Study , if a subject receives 2 different valve types from 2 different manufacturers, the subject will be excluded from the implanted analysis.

	Among the roll-in, ACURATE *Prime* XL, and Continued Access Study cohorts, for the ITT analysis, all subjects who sign an Informed Consent Form and are enrolled in the study will be included in the analysis sample, regardless of whether the test device was implanted. The Implanted analysis set will include all subjects who sign an Informed Consent Form and are implanted with an ACURATE valve. For these single-arm cohorts, the subject is considered enrolled in the trial when there is an attempt made to insert the ACURATE Transfemoral Aortic Valve System into the subject's femoral artery.
Primary Endpoint Statistical Hypothesis	In the Main Randomized Cohort, the primary endpoint (composite of all-cause mortality, all stroke, and rehospitalization at 1 year) rate for the ACURATE group is non-inferior to that for the Control group.
Statistical Test Method for the Primary Endpoint	For the Main Randomized cohort, the statistical hypothesis is that the primary endpoint rate for ACURATE is non-inferior to the rate for Control: H_0: $P_{E\_ACURATE}$ minus $P_{E\_Control} \geq \Delta$ (Inferior) H_1: $P_{E\_ACURATE}$ minus $P_{E\_Control} < \Delta$ (Non-inferior) where $P_{E\_ACURATE}$ and $P_{E\_Control}$ correspond to the primary endpoint rates for the ACURATE group (test) and the Control group, respectively, and Δ is the non-inferiority margin. The primary analysis set for the primary endpoint is the ITT analysis set. This endpoint will also be analyzed for the implanted analysis set. A Bayesian analysis[d] will be performed to estimate the treatment difference between ACURATE and Control through posterior probability. d: Popma JJ, et al. N Engl J Med 2019;380:1706-15 Reardon MJ, et al. N Engl J Med 2017;376:1321-31
Sample Size Parameters for the Primary Endpoint	Although the primary endpoint analysis is performed using the Bayesian method, the sample size calculation for the Main Randomized Cohort (see **Note 12** below) is based on a standard non-inferiority two-sample test approach. The sample size calculation for the primary endpoint is based on the following assumptions. • Expected rate for both arms = 22.3% (based on weighted average of TAVR data[e] see **Note 13** below) • Non-inferiority margin (Δ) =8.0% (36% relative to the expected rate) • Test significance level (α) = 0.025 (1-sided) (see **Note 14** below)

	• Test (ACURATE): Control ratio = 1:1 • Power (1 minus ß) > 90% • Expected rate of attrition = 5% • Total sample size = 1500 (750 per group) • Number of evaluable subjects per group = 712 • Analyses: One administrative, one formal interim, one final (see **Note 15** below) **Note 12:** The Pocock-type method[f] is used during sample size calculations. The statistical software EAST® 6.5 is used for the sample size calculations. **Note 13:** The estimated proportions of subjects by operative risk level is 10% extreme risk, 25% high risk, 30% intermediate risk, and 35% low risk. **Note 14:** A statistically equivalent posterior probability threshold for the Bayesian analysis is empirically chosen through extensive simulations. **Note 15:** The administrative interim analysis will be conducted when the first 350 subjects in the Main Randomized Cohort have completed 1-year follow-up. The formal interim analysis will be carried out after enrollment in the Main Randomized Cohort is completed. This formal interim analysis will be conducted on the full N=1500 subject cohort of the Main Randomized Cohort after first 1050 subjects in Main Randomized Cohort have completed 1-year follow-up. The piecewise exponential model[g] based on outcomes among these subjects will be used to estimate the 1-year results by treatment group. The posterior distributions for the parameters of interest will be used to evaluate the hypothesis testing. A final analysis for the primary endpoint will be performed on all subjects with completed 1-year data if non-inferiority cannot be claimed at the formal interim analysis (see Success Criteria below). e: Adams DH, et al. N Engl J Med 2014;370:1790-8 Mack MJ, et al. N Engl J Med 2019;380:1695-705 Popma JJ, et al. J Am Coll Cardiol 2014;63:1972-81 Popma JJ, et al. N Engl J Med 2019;380:1706-15 Reardon MJ, et al. N Engl J Med 2017;376:1321-31 Smith CR, et al. N Engl J Med 2011;364:2187-98 Thourani VH, et al. Lancet 2016;387:2218-25 Waksman R, et al. J Am Coll Cardiol Intv 2019;12:901–7 Webb JG, et al. J Am Coll Cardiol Intv 2015;8:1797-806 Medtronic CoreValve System PMA P130021/S002: FDA Summary of Safety and Effectiveness Data f: Lan KKG and DeMets DL. Biometrika 1983;70:659–63 g: Popma JJ, et al. N Engl J Med 2019;380:1706-15

	Note: Addition to the criterion of the 1-year completion of the first 1050 subjects in the Main Randomized Cohort, the formal interim analysis will be conducted only after at least 80% low risk subjects in Main Randomized Cohort have completed the 6-month follow-up, whichever comes later.
Success Criteria for the Primary Endpoint	The Bayesian method is used to test the non-inferiority hypothesis of the primary endpoint on the Main Randomized Cohort. To establish that the ACURATE device is non-inferior to the Control, the results will need to meet the following equation: $Pr(H_1 \mid Data) > \xi$ where • $Pr(H_1 \mid Data)$ is the posterior probability of H_1 given the observed data at either the interim or the final analysis; • H_1 is the alternative hypothesis for non-inferiority: PE_ACURATE minus PE_Control < Δ; • ξ is a prespecified threshold, which is empirically chosen through extensive simulations using the Bayesian approach for the noninferiority tests. If non-inferiority has been declared at the formal interim analysis, the non-inferiority test will not be performed at the final analysis. If noninferiority cannot be declared at this interim analysis, the non-inferiority test will be performed at the final analysis for all subjects using the Bayesian method with the same pre-specified threshold. The study will not stop for futility at the interim analysis.
ACURATE *Prime* XL Nested Registry Statistical Hypothesis	Mean aortic valve pressure gradient at 30 days post implant procedure is less than a performance goal (PG): H_0: $Gradient_{30D} \geq PG$ H_1: $Gradient_{30D} < PG$ where $Gradient_{30D}$ is the 30-day mean aortic valve pressure gradient for the ACURATE *Prime* XL valve and PG is 15 mmHg.
Statistical Test Method for the ACURATE *Prime* XL Nested Registry	A one-sample *t*-test will be used to test the one-sided hypothesis at a significance level of 2.5%.
Sample Size Parameters for the ACURATE	• Expected 30-day mean pressure gradient from ACURATE *Prime* XL = 10 mmHg[h] • Expected standard deviation = 7 mmHg

Prime XL Nested Registry	• PG = 15 mmHg • Test significance level (α) = 0.025 (1-sided) • Power > 90% • Evaluable number of subjects = Minimum of 40 subjects • Expected rate of attrition = 20% (8 subjects) • Planned enrollment of 50 subjects • The analysis population for the hypothesis testing will be the subject population implanted with the ACURATE *Prime* XL valve. h: Based on Boston Scientific data on file and published data for large-annulus CoreValve devices: Tang GHL, et al. Am J Cardiol 2019;124:1091-8 Kalogeras K, et al. Catheter Cardiovasc Interv 2019;93:685-91 Ussia GP, et al. EuroIntervention 2015;10:e1-e8
Success Criteria for the ACURATE *Prime* XL Nested Registry	If the *P* value from the one-sample *t*-test is < 0.025, the ACURATE *Prime* XL valve will be concluded to have a 30-day mean aortic valve pressure gradient < 15 mmHg. This corresponds to the one-sided upper 2.5% confidence bound of the observed 30-day mean aortic valve pressure gradient being < 15 mmHg.
Statistics Summary	Data will be summarized separately for each cohort. In addition to the hypothesis tests listed above, descriptive statistics will be used. Full methods will be described in the statistical analysis plan.

2 INTRODUCTION

The study is designed to assess the safety and effectiveness of the ACURATE Transfemoral Aortic Valve System for transcatheter aortic valve replacement (TAVR) in subjects with severe native aortic stenosis who are indicated for TAVR. This analysis plan specifies the detailed statistical methods and definitions used for the endpoint analyses and statistical data reporting.

3 ENDPOINT ANALYSIS

The Main Randomized Cohort data from the ACURATE IDE trial will be used for the primary endpoint analysis (including 4D CT subjects). Data will also be summarized separately for the randomized subjects in the 4D CT Imaging Substudy.

Data will be summarized separately for subjects in the ACURATE IDE roll-in cohort , Extended Durability Cohort, and Continued Access Study cohorts. Descriptive statistics will be used to summarize the data in these cohorts and no statistical inference will be made.

Data will be summarized separately for the specific statistical hypothesis associated with the ACURATE *Prime*™ XL Nested Registry. Descriptive statistics also will be used to summarize data from subjects in the single-arm ACURATE *Prime* XL cohort.

3.1 Primary Endpoint for the Main Randomized Cohort

The primary endpoint for the Main Randomized Cohort is the composite of all-cause mortality, all stroke, and rehospitalization† at 1 year. The events in this endpoint will be adjudicated by an independent CEC.

† Hospitalization for valve-related symptoms or worsening congestive heart failure (NYHA class III or IV); per VARC-2 definition.

3.1.1 Hypotheses

The statistical hypothesis is that the primary endpoint (composite of all-cause mortality, all stroke, and rehospitalization at 1 year) rate for the ACURATE (test) group is non-inferior to that for the Control group.

The null and alternative hypotheses for the primary endpoint are as follows:

H_0: $P_{E\_ACURATE}$ minus $P_{E\_Control} \geq \Delta$ (Inferior)

H_1: $P_{E\_ACURATE}$ minus $P_{E\_Control} < \Delta$ (Non-inferior)

where $P_{E\_ACURATE}$ and $P_{E\_Control}$ correspond to the primary endpoint rates for the ACURATE and the Control group, respectively, and Δ is the non-inferiority margin.

The primary analysis set for the primary endpoint is the ITT analysis set. This endpoint will also be analyzed for the implanted analysis set.

A Bayesian analysis will be performed to estimate the treatment difference between ACURATE and Control through posterior probability.

3.1.2 Sample Size

Although the primary endpoint analysis is performed using the Bayesian method, the sample size calculation for the Main Randomized Cohort is based on a standard non-inferiority two-sample test approach with one formal interim analysis planned. The sample size calculation for the primary endpoint is based on the following assumptions.

- Expected rate for both arms = 22.3% (based on a weighted average of TAVR data
- Non-inferiority margin (Δ) = 8.0% (36% relative to expected rate)
- Test significance level (α) = 0.025 (1-sided)
- Test (ACURATE): Control ratio = 1:1
- Power (1 minus ß) > 90%
- Expected rate of attrition = 5%
- Total sample size = 1500 (750 per group)
- Number of evaluable subjects per group = 712
- Analyses: One administrative, one formal interim, one final

The Pocock-type method is used during sample size calculations when 70% subjects reached 1-year follow-up at the formal interim analysis. The statistical software EAST® 6.5 is used for the sample size calculations.

The estimated proportions of subjects in the Main Randomized Cohort by operative risk level are 10% extreme risk, 25% high risk, 30% intermediate risk, and 35% low risk.

The formal interim analysis will be carried out after enrollment in the Main Randomized Cohort is completed. This formal interim analysis will be conducted on the full N=1500 subjects of the Main Randomized Cohort after first 1050 subjects in the Main Randomized Cohort have completed 1-year follow-up or at least 80% low risk subjects in the Main Randomized Cohort have completed 6-month follow-up, whichever comes later. The piecewise exponential models will be based on outcomes among these subjects who have completed 30-day, 6-month, or 1-year follow-ups. The results from the models will be used to estimate the 1-year results by treatment group through Bayesian posterior probabilities for the non-inferiority test. The final analysis for the primary endpoint will be performed on all subjects with completed 1-year data if non-inferiority cannot be claimed at the formal interim analysis.

3.1.3 Success Criteria

The Bayesian method is used to test the non-inferiority hypothesis of the primary endpoint. To establish that the ACURATE device is non-inferior to the Control, the results will need to meet the following equation:

$$Pr(H_1 \mid Data) > \xi$$

where

- $Pr(H_1 \mid Data)$ is the posterior probability of H_1 given the observed data at either the interim or the final analysis;
- H_1 is the alternative hypothesis for non-inferiority: $PE_{\_ACURATE}$ minus $PE_{\_Control} < \Delta$;
- ξ is a prespecified threshold, which is empirically chosen through extensive simulations using the Bayesian approach for the noninferiority tests.

If non-inferiority has been declared at the formal interim analysis, the non-inferiority test will not be performed at the final analysis. If non-inferiority cannot be declared at this interim analysis, the non-inferiority test will be performed at the final analysis for all subjects using the Bayesian method with the same pre-specified threshold. The study will not stop for futility at the interim analysis.

3.1.4 Bayesian Statistical Methods

3.1.4.1 Piecewise exponential survival model

The primary endpoint event rates for the hypothesis test specified in Section 3.1.1 will be estimated through posterior probabilities of the hazards from the piecewise exponential survival models. The piecewise exponential models will be constructed by subject operative risk (extreme/high, intermediate, and low). To simplify the notations, the formulas in the Section 3.1.4 represent the general notations within each risk stratum and the same notations can be applied to the data for each risk stratum.

Within each risk stratum, the time partition for the piecewise exponential models between 0-365 days is defined as:

$$0 = t_0 < t_1 < t_2 < t_3 = 365, \text{ where } t_1 = 30 \text{ and } t_2 = 180$$

The hazard rates will be estimated for the three intervals: $\tau_1: [t_0, t_1]$, $\tau_2: (t_1, t_2]$, and $\tau_3: (t_2, t_3]$. The hazard within each time interval is considered as constant by treatment group.

Within each time intervals, the piecewise exponential models will be generated by treatment group (test and control). Let $\lambda_{j,k}$ denote the hazard rate for subjects in the j[th] treatment group and the k[th] time interval, j=1, 2 and k=1, 2, 3, where the indices are defined as the following:

j is for treatment groups (1: test arm, 2: control arm),
k is for the time intervals (1: [0,30], 2: (30, 180], 3: (180, 365])

There will be total six estimated the hazard rates within each risk stratum.

Based on the estimated hazard rates $\lambda_{j,k}$ from the models, the estimated probability of a subject who is in the j[th] treatment group developing an event at time $\tilde{t}$ is defined as:

$$P_j(\tilde{t}; \lambda_{j,k}) = \begin{cases} 1 - e^{-\left(\lambda_{j,1}*(\tilde{t}-t_0)\right)}, & t_0 \le \tilde{t} \le t_1 \\ 1 - e^{-\left(\lambda_{j,1}*(t_1-t_0)+\lambda_{j,2}*(\tilde{t}-t_1)\right)}, & t_1 < \tilde{t} \le t_2 \\ 1 - e^{-\left(\lambda_{j,1}*(t_1-t_0)+\lambda_{j,2}*(t_2-t_1)+\lambda_{j,3}*(\tilde{t}-t_2)\right)}, & t_2 < \tilde{t} \le t_3 \end{cases}$$

Where $\lambda_{j,k} = (\lambda_{j,1}, \lambda_{j,2}, \lambda_{j,3})$.

3.1.4.2 Prior distribution

For the piecewise exponential models, the Gamma distribution is used as the prior distribution $\lambda_{j,k} \sim Gamma(a, b)$ for the hazard parameters. The Gamma distribution $Gamma(a, b)$ has the following PDF:

$$f_{a,b}(\lambda_{j,k}) = b^a {\lambda_{j,k}}^{a-1} e^{-b\lambda_{j,k}} / \Gamma(a), a, b > 0, \lambda_{j,k} > 0$$

where $\Gamma(a)$is a Gamma function, a is the shape parameter and b is the inverse-scale parameter.

The Gamma distribution for the prior uses the shape parameter $a = 0.0001$ and the inverse-scale $b = 0.0001$. The mean of the prior distribution is equal to 1 with a variance of 10000. The independent Gamma prior will be used for all parameters $\lambda_{j,k}$. This prior is proper and reasonably noninformative.

3.1.4.3 Posterior probability

Let λ_j denotes the vector of parameters $(\lambda_{j,1}, \lambda_{j,2}, \lambda_{j,3})$ for the hazard rates from the model of j[th] treatment group. $L(D_j|\lambda_j)$ is the likelihood function defined as:

$$L(D_j|\lambda_j) = \prod_{k=1}^{3} (\lambda_{j,k})^{d_k} \times e^{-\lambda_{j,k}*T_k}$$

$$d_k = \sum_{m=1}^{n} I_{\tau_k}(\tilde{t}_m)$$

$$T_k = \sum_{m=1}^{n} J_{\tau_k}(\tilde{t}_m)$$

$$I_{\tau_k}(\tilde{t}_m) = \begin{cases} 1, if\ event\ occurred\ at\ \tilde{t}_m \in \tau_k \\ 0, if\ no\ event\ occurred\ up\ to\ \tilde{t}_m \in \tau_k \end{cases}$$

$$J_{\tau_k}(\tilde{t}_m) = \begin{cases} \tilde{t}_m - t_{k-1}, if\ t_{k-1} \le \tilde{t}_m < t_k \\ t_k - t_{k-1}, if\ \tilde{t}_m \ge t_k \\ 0, \quad if\ \tilde{t}_m < t_{k-1} \end{cases}$$

where $0 \le \tilde{t}_m \le 365$, m=1, …, n. $\tilde{t}_m$ is the number of days that the m[th] subject has been in the study. The n is the number of subjects in D_j. The D_j is the observed data for the j[th] treatment arm to fit the model. τ_1: $[t_0, t_1]$, τ_2: $(t_1, t_2]$, and τ_3: $(t_2, t_3]$, where $t_0 = 0,\ t_1 = 30$, $t_2 = 180$, and $t_3 = 365$.

The posterior distribution is proportional to a joined distribution of the prior and likelihood function that can be constructed as following:

$$\Pr(\lambda_j | D_j) \propto f_{a,b}(\lambda_j) * L(D_j | \lambda_j) \propto \lambda_j^{(a+d_k)-1} e^{-(b+T_k)\lambda_j}$$

$\Pr(\lambda_j | D_j)$ follows a gamma distribution that the shape parameter equals to $(a + d_k)$ and the inverse-scale parameter equals to $(b + T_k)$, where k=1, 2, 3 and $a = 0.0001, b = 0.0001$.

The posterior distribution for each lambda is in the closed form of a Gamma distribution. Deriving the full conditional posterior distributions of the primary endpoint can be mathematically challenging. As stated in User's Guide Introduction to Bayesian Analysis Procedure: "When models become too difficult to analyze analytically, you have to use simulation algorithms, such as the Markov chain Monte Carlo (MCMC) method to obtain posterior estimates". Bayesian piecewise exponential model method will use the Markov chain Monte Carlo (MCMC) method to simulate posterior distributions of all lambdas. The posterior distributions of the primary endpoint then can be constructed based on the posterior distributions of the lambdas and the prespecified piecewise exponential models. The statistical inference for the primary endpoint will be based on the simulated posterior distributions of the lambdas and the constructed posterior distributions of the primary endpoint. The sampling algorithm for the posterior distributions implemented in SAS could be found in the SAS Proc MCMC procedure. The posterior distribution of the lambda for each interval will be independently constructed using the Metropolis-Hastings algorithm based on the likelihood function and a time-to-event data for a specific time interval.

Reference:
SAS/STAT®14.1 User's Guide Introduction to Bayesian Analysis Procedure, SAS Institute Inc.
SAS Manual, https://support.sas.com/documentation/onlinedoc/stat/131/mcmc.pdf

3.1.5 Interim Analysis

3.1.5.1 Completed visits at 30-day, 6-month, and 1-year

The interim analysis for the primary endpoint will be conducted after 1050 subjects in the Main Randomized Cohort have completed a 1-year follow-up, or at least 80% low risk subjects in the Main Randomized Cohort have completed a 6-month follow-up, whichever comes later. A subject is considered as completion of 1-year follow-up if the subject meets the following by the pre-specified cutoff date (defined in 3.1.5.2):

1) has completed 1-year follow-up visit, or
2) has not completed 1-year follow-up visit and is reported in the 1-year visit form as 'not done' or in the End-of-Study form (as died, withdrew, or lost to 1-year follow-up) and has reached 365 days post index procedure (or randomization if no procedure), or
3) has not completed 1-year follow-up visit and is not reported as 'not done' or not in the End-of-Study (as died, withdrew, or lost to 1-year follow-up) and has reached 395 days post index procedure (or randomization if no procedure).

Similarly, a subject is considered as completed 30-day and 6-month follow-up visits as the following:

For 30-day visit by the pre-specified cutoff date (defined in 3.1.5.2):

1) has completed 30-day follow-up visit, or
2) has not completed 30-day follow-up visit and is reported in the 30-day visit form as 'not done' or in the End-of-Study form (as died, withdrew, or lost to 30-day follow-up) and has reached 30 days post index procedure (or randomization if no procedure), or
3) has not completed 30-day follow-up visit and is not reported as 'not done' or not in the End-of-Study (as died, withdrew, or lost to 30-day follow-up) and has reached 37 days post index procedure (or randomization if no procedure).

For 6-month visit by the pre-specified cutoff date (defined in 3.1.5.2):

1) has completed 6-month follow-up visit, or
2) has not completed 6-month follow-up visit and is reported in the 6-month visit form as 'not done' or in the End-of-Study form (as died, withdrew, or lost to 6-month follow-up) and has reached 180 days post index procedure (or randomization if no procedure), or
3) has not completed 6-month follow-up visit and is not reported as 'not done' or not in the End-of-Study (as died, withdrew, or lost to 6-month follow-up) and has reached 210 days post index procedure (or randomization if no procedure).

3.1.5.2 Formal interim analysis cutoff date and analysis cohorts

A cutoff date for the interim analysis is the date when the first 1050 subjects in the Main Randomized Cohort have completed a 1-year follow-up or at least 80% low risk subjects in the Main Randomized Cohort have completed a 6-month follow-up, whichever comes later. All subjects who have completed 30-day, 6-month, or 1-year (defined in Section 3.1.5.1) prior to or on the cutoff date should be included in the 30-day, 6-month, or 1-

year analysis cohort, respectively. The three analysis cohorts will be used for the piecewise exponential models for different time intervals, defined as the following:

- 30-day analysis cohort will be used for piecewise exponential models between 0-30 days.
- 6-month analysis cohort will be used for piecewise exponential models between 31-180 days.
- 1-year analysis cohort will be used for piecewise exponential models between 181-365 days.

Note: the analysis cohorts (30-day, 6-month, 1-year) are for ITT analysis. The analysis cohorts for the Implanted analysis will the subset of the ITT set for subjects who meet the Implanted definition.

3.1.5.3 Models for estimating parameters

The piecewise exponential survival model will be constructed separately for each risk stratum (extreme/high, intermediate, and low), each treatment group (test and control), and each time interval (30-day, 6-month, and 1-year) on subjects who have completed respective visits. There will be total eighteen models (nine models within each treatment group). Each model will estimate one parameter of the lambdas and simulate the corresponding posterior distribution as shown in (Table 1) of the $\lambda_{i,j,k}$ (i=1, 2 for test and control groups; j=1, 2, 3 for E/HR, IR, LR; k=1, 2, 3 for the three time intervals):

Table 1. Lambdas by Risk Group, Time Interval, and Treatment

Risk Strata/Time Intervals	**Patients completed 30 days [0-30] days**	**Patients completed 180 days [31-180] days**	**Patients completed 365 days [181-365] days**
	Estimated Lambdas (Test Group)		
Extreme or high risk (E/HR)	$\lambda_{1,1,1}$	$\lambda_{1,1,2}$	$\lambda_{1,1,3}$
Intermediate risk (IR)	$\lambda_{1,2,1}$	$\lambda_{1,2,2}$	$\lambda_{1,2,3}$
Low risk (LR)	$\lambda_{1,3,1}$	$\lambda_{1,3,2}$	$\lambda_{1,3,3}$
	Estimated Lambdas (Control Group)		
Extreme or high risk (E/HR)	$\lambda_{2,1,1}$	$\lambda_{2,1,2}$	$\lambda_{2,1,3}$
Intermediate risk (IR)	$\lambda_{2,2,1}$	$\lambda_{2,2,2}$	$\lambda_{2,2,3}$
Low risk (LR)	$\lambda_{2,3,1}$	$\lambda_{2,3,2}$	$\lambda_{2,3,3}$

3.1.5.4 Bayesian method for the hypothesis test

For each Bayesian piecewise exponential model, 10000 posterior samples for each $\lambda_{i,j,k}$ are generated, the estimated 1-year primary endpoint event rates ($PE_{\_ACURATE}$ and $PE_{\_Control}$) for each sample of $\lambda_{i,j,k}$ can be calculated based on the following:

$$PE_{\_ACURATE,\ E/HR} = 1 - e^{-(31*\lambda_{1,1,1}+150*\lambda_{1,1,2}+185*\lambda_{1,1,3})},$$
$$PE_{\_Control,\ E/HR} = 1 - e^{-(31*\lambda_{2,1,1}+150*\lambda_{2,1,2}+185*\lambda_{2,1,3})},$$
$$PE_{\_ACURATE,\ IR} = 1 - e^{-(31*\lambda_{1,2,1}+150*\lambda_{1,2,2}+185*\lambda_{1,2,3})},$$
$$PE_{\_Control,\ IR} = 1 - e^{-(31*\lambda_{2,2,1}+150*\lambda_{2,2,2}+185*\lambda_{2,2,3})},$$
$$PE_{\_ACURATE,\ LR} = 1 - e^{-(31*\lambda_{1,3,1}+150*\lambda_{1,3,2}+185*\lambda_{1,3,3})},$$
$$PE_{\_Control,\ LR} = 1 - e^{-(31*\lambda_{2,3,1}+150*\lambda_{2,3,2}+185*\lambda_{2,3,3})},$$

The overall weighted 1-year primary endpoint rates for each sample from the posterior distributions are calculated as the following:

$$PE_{\_ACURATE} = 35\%* PE_{\_ACURATE,\ E/HR} + 30\%* PE_{\_ACURATE,\ IR} + 35\%* PE_{\_ACURATE,\ LR}$$
$$PE_{\_Control} = 35\%* PE_{\_Control,\ E/HR} + 30\%* PE_{\_Control,\ IR} + 35\%* PE_{\_Control,\ LR}$$

The treatment difference ($PE_{\_ACURATE}$ minus $PE_{\_Control}$) will be calculated. Iterating this sampling process for 10000 times over the posterior distributions of the lambdas, the posterior distributions of the primary endpoint rate by treatment group and the posterior distribution of the treatment difference can be constructed for the hypothesis test on the primary endpoint. Similarly, the credible intervals will be constructed based on these posterior distributions. The non-inforiority test then is performed based on the pre-specified success criteria: the posterior probability of H_1 given the observed interim data is greater than the threshold ξ. If the posterior probability is greater than the pre-specified threshold ξ, then the non-inferiority is declared. To be specific, if the posterior probability of the treatment difference ($PE_{\_ACURATE}$ minus $PE_{\_Control}$) < Δ constructed from the sampling distribution (N=10000) is greater than the pre-specified threshold ξ, then the non-inferiority is declared.

Reference:
Wilber DJ, et al. *JAMA. 2010;303(4):333-340.*
Popma JJ, et al. N Engl J Med 2019;380:1706-15
Reardon MJ, et al. N Engl J Med 2017;376:1321-31

3.1.6 Final Analysis

If non-inferiority has been declared at the formal interim analysis, the non-inferiority test will not be performed at the final analysis. Otherwise, the non-inferiority test will be performed at the final analysis using Bayesian method when all subjects have completed 1-year data. As it has been formulized in the Section 3.1.3, the same piecewise exponential model and pre-specified success criteria are used. If the posterior probability of H_1 given the observed data is greater than pre-specified threshold ξ, then the non-inferiority is declared at the final analysis for the primary endpoint.

3.1.7 Sensitivity Analyses for Primary Endpoint

Handling of dropouts and missing data will depend on their frequency and the nature of the outcome measure. Sensitivity analyses (e.g., tipping-point analysis) will be performed

to assess the impact of subjects with inadequate follow-up (i.e., missing data) on the primary endpoint and to assess the robustness of the conclusion of the primary analysis. The tipping-point analysis is based on the exact same data set used in the main primary endpoint analysis. A patient is considered as having missing primary endpoint data when the patient's last day in the study is less than 335 days post-randomization without occurrence of any primary endpoint events, e.g.: patients withdrew, lost to follow up, or had 1-year visit prior to reaching 335 days post-randomization and had no primary endpoint events. In the tipping point analysis, these patients will be considered as not having sufficient follow-up days. The tipping-point analysis will impute the primary endpoint event occurred one day post censoring (days to imputed primary endpoint event) for the subject with not sufficient 1-year follow-up. The missing cases will be imputed by the chronological order of the imputed days to the primary endpoint event and then by the randomization date or procedure date. The tipping point analysis assumes the subject with missing data will develop a primary endpoint event at the time of the last follow-up days (<335 days) for the test arm, but not develop any primary endpoint event for the control arm. The tipping point analysis results will show how many more subjects on average in the test arm could develop a primary endpoint event while the study could still declare the non-inferiority. The tipping point table and plot will be provided to demonstrate the robustness of the conclusion of the primary endpoint hypothesis test. Tipping point analysis will use the same Bayesian approach. Tipping point analysis will be provided only when the non-inferiority is declared at either the interim or the final.

The time intervals of the main piecewise exponential model are based on the actual study follow-up visits defined in the protocol (such as: 30-day, 6-month, and 1-year). The time interval sensitivity analyses will be based on the exact same data set that the main primary analysis will be conducted on. For this sensitivity analysis, each time interval (such as: 0-30 days, 31-180 days, and 181-365 days) will be partitioned into two intervals with approximately equal number of events. The Proc PHREG procedure will be used. The additional sensitivity analysis using different time intervals between 0-365 days (eg: the SAS default from the Proc PHREG procedure that partitions the time axis into eight intervals with approximately equal number of events in each interval when appropriate) for the piecewise exponential model may be performed for the primary endpoint analysis. This sensitivity analysis will only be performed on all the patients who completed the 1-year visits. For each individual model, if the model with eight intervals by default doesn't exist, number of intervals will be reduced to four intervals. If model with four intervals still doesn't exist, the number of intervals for the model will be reduced by one each time until a valid model is identified. Partitions and number of time intervals between different models could be different.

In the main analysis using Bayesian approach, subjects are considered exchangeable between the earlier enrolled and the later enrolled. Thus, the assumption is that the subject characteristics are distributed at random over the enrollment time. In order to check the model robustness, sensitivity analysis for the primary endpoint may be performed with piecewise exponential model adjusting for key covariates, such as: control devices for the control arm only (CoreValve or SAPIEN 3).

The sensitivity analysis for the control devices will fit three piecewise exponential models based on the following groups of subjects.

1. Test group subjects by risk strata and time intervals
2. CoreValve subjects by risk strata and time intervals (N=w1 subjects)
3. SAPIEN 3 subjects by risk strata and time intervals (N=w2 subjects)

Model 2 and 3 results will be combined using the weighted method to estimate the primary endpoint rate. The weights are proportional to the numbers of observations between the control devices (CoreValve vs SAPIEN) from the observed data that are used to fit each model. The weights are w1/(w1+w2) and w2/(w1+w2) for Model 2 and Model 3, respectively. With the sampling of posterior distributions of the hazards, the hypothesis can be tested.

Sensitivity analysis may be performed using the observed weights for the operative risk cohorts. The weights are proportional to the numbers of observations between the operative risks (extreme/high, intermediate, and low) from the observed data that are used to fit each model. The sensitivity analysis is based on the exact same data set used in the main primary endpoint analysis.

Note: the sensitivity analyses, tipping point analyses, or analyses involved the piecewise exponential modes may be adjusted based on the pre-specified methods (e.g. if pre-specified the number of events are too small for a piecewise exponential model, which causes the model not converged. The number of intervals or number of models or other related parameters may be reduced or adjusted accordingly.)

Reference: SAS Manual,
https://support.sas.com/documentation/onlinedoc/stat/131/phreg.pdf

3.1.8 Bayesian Analysis Results

For the Bayesian analysis, the following results will be provided as appropriate:

- the posterior probability for non-inferiority
- the posterior medians by treatment arm and their 95% credible intervals
- the posterior median of treatment difference ($PE_{\_ACURATE}$ minus $PE_{\_Control}$) and it's 95% credible intervals
- the highest posterior density (HPD) intervals and the equal-tail posterior intervals for the estimated parameters
- other data derived from the posterior distribution, such as: model fitting statistics, posterior mean, and graphic presentation of the posterior distributions of the parameters

3.1.9 Bayesian Design Operating Characteristics

3.1.9.1 Simulation setup

<u>Piecewise exponential model:</u>

1) Time intervals follow the actual follow-up visits defined in the protocol, which are the intervals as 0 - 30 days, 31 – 180 days, and 181 – 365 days post randomization for the ITT analysis.

2) Hazard rate within each time interval is constant. The hazard rates may be different between different intervals. The following is the hazard function defined by different treatment groups:

$$h_j(t) = \begin{cases} \lambda_{j,1}, & 0 \le t \le 30\ days \\ \lambda_{j,2}, & 31 \le t \le 180\ days \\ \lambda_{j,3}, & 181 \le t \le 365\ days \end{cases}$$

Where j=1 or 2 for the test and the control groups within each risk stratum, respectively.

3) SAS Programming Software and the Procedure PROC MCMC are used for simulation and the Bayesian analyses.

Note: the piecewise exponential models are constructed by risk stratum, treatment group, and time interval. The parameters to be estimated are specified in Section 3.1.5.2 Table 1.

Simulation parameters

1) Proper and reasonably non-informative prior Gamma (shape=0.0001, inverse-scale=0.0001)
2) Assumption of number of subjects available for each time interval at the time of the interim analysis by risk stratum for each treatment group (total N=750 per treatment group) (Table 2).

Table 2. Available Data for Simulations at Interim for each treatment group

Risk Stratum	30-Day	6-Month	1-Year
E/HR	262	241	194
IR	225	225	221
LR	262	223	115

3) Expected rate for the Control (22.3%)
4) Expected rate for the Test (30.3% for alpha, 22.3% for the power).
5) Hazard rates assumptions by the time intervals as following (Table 3):

Table 3. Cases for the Relationship between Lambdas for Simulations

Hazard assumptions*	λ2	λ3
Case 1	0.4*λ1	0.3*λ1
Case 2	0.3*λ1	0.2*λ1
Case 3	0.3*λ1	0.1*λ1
Case 4	0.2*λ1	0.2*λ1

*The assumptions are estimated based on historical data for event rates at 30-day, 6-month, and 1-year (Table 4).

Table 4. Historical Event Rate (Observed or Estimated) by Time Intervals

Clinical Study	Device	Risk Profile	Available Event	Event rate by time 30d vs 6m vs 12m	Estimated Lambdas (λ1, λ2, λ3)
Partner 2B	SAPIEN XT	Extreme	death/major stroke/rehosp	17% vs 28% vs 37%	λ2=0.15*λ1 λ3=0.12*λ1
CoreValve ER	Corevalve	Extreme	death/major stroke	10% vs 19% vs 26%	λ2=0.19*λ1 λ3=0.14*λ1
Partner 1A	SAPIEN	High	death/major stroke	7% vs 18% vs 26%	λ2=0.33*λ1 λ3=0.23*λ1
CoreValve HR	CoreValve	High	death/major stroke	5.2% vs 11.5% vs 16.3%	λ2=0.25*λ1 λ3=0.17*λ1
CoreValve IR	CoreValve	Intermediate	death/major stroke	2.8% vs 6.0% vs 8.1%	λ2=0.23*λ1 λ3=0.13*λ1
Evolut LR	CV 31mm, Evolut R/Pro	Low	death/major stroke	0.8% vs 1.9% vs 2.9%	λ2=0.27*λ1 λ3=0.21*λ1
PARTNER 3	Sapien 3	Low	death/stroke/rehosp	4.2% vs 6.5% vs 8.5%	λ2=0.11*λ1 λ3=0.08*λ1
					λ1>λ2>λ3
				Range for λ2 related to λ1 Mean [min, max]	0.22 [0.11, 0.33] of λ1
				Range for λ3 related to λ1 Mean [min, max]	0.15 [0.08, 0.23] of λ1

6) Run 2000 trials and draw 3,000 samples for each trial.
7) Seeds used for piecewise exponential models is 12345.
8) Sensitivity analysis by time intervals.
   The purpose for the sensitivity analyses is for evaluate the robustness of the prespecified piecewise exponential models in terms of type I errors.
   The following time interval will be used for constructing the sample data for simulations
   Case 1: [0, 7], [8, 30], [31, 365]
   Case 2: [0, 7], [8, 90], [91, 365]
   Case 3: [0, 90], [91, 180], [181, 365]
   Case 4: [0, 180], [181, 270], [271, 365]
   For each case, the sample data distribution will follow these time intervals with constant hazard rate within each interval. However, the analysis using the piecewise exponential model will follow the prespecified time intervals ([0, 30], [31, 180], [181, 365]) for the hypothesis test of the primary endpoint. The sensitivity analysis for type I error will be evaluated based on the piecewise exponential models with pre-specified fixed time intervals ([0, 30], [31, 180], [181, 365]) and on the sample data that are not following the same distribution (for example in Case 1: data with constant hazard rates within [0, 7], [8, 30], [31, 365]).

3.1.9.2 Summary of Simulation Results

The type I error is simulated when the observed Test rate minus the observed Control rate is equal to the prespecified margin 8.0% while the observed Control rate is as the expected 22.3%. The power is simulated when the observed Test Rate is equal to the observed Control rate as the expected 22.3%.

After considering all simulation results from different simulation parameters specified in Section 3.1.9, we choose $\xi = 97.5\%$ as the threshold for the posterior probability threshold for testing treatment difference of $PE\_{ACURATE}$ minus $PE\_{Control}$ for the non-inferiority test. Given this threshold, the simulations show that the type I error is reasonably controlled at 2.5% for the interim and final analyses. While in a few cases, the type I errors are slightly greater than 2.5% due to the simulation random error and the total number of models (the eighteen models for each trials), they are well within the Exact 95% confidence interval [1.9%, 3.3%] of the desired 2.5% Type I error for the 2000 samples. For the expected hazards estimated from other studies when $\lambda2=0.3*\lambda1$ and $\lambda3=0.2*\lambda1$, the type I errors are less than 2.5% and within the confidence interval of the 2.5% for both the interim and the final. As the results show, the type I error results are consistent across different scenarios. No evidence shows that these numbers are statistical different from the 2.5%, thus Type I error is reasonably controlled (Table 5). Considering variabilities from time to event sample data and each piecewise exponential model (from 18 models) with the process of the Bayesian posterior sampling for all the parameters, and combining with robustness of the simulation results, it is believed that piecewise exponential models will model the data well. The selection of the threshold for the hypothesis test is reasonable and will provide a robust evaluation on the primary endpoint result for the study.

Table 5. Simulation Results for Type I Errors Simulated type I error

Simulation Parameters				**Chance of the Posterior Probability for Test Rate minus Control Rate <8% is greater than ξ**				
Test Rate	**Control Rate**	**Hazard Rate**	**Time Interval***	**Time**	**ξ=97.1%**	**ξ=97.3%**	**ξ=97.5%**	**ξ=97.7%**
30.30%	22.3%	λ2=0.4*λ1 λ3=0.3*λ1	[0, 30] [31, 180] [181,365]	Interim	2.70%	2.55%	2.45%	2.30%
				Final	0.75%	0.65%	0.55%	0.60%
30.30%	22.3%	λ2=0.3*λ1 λ3=0.2*λ1	[0, 30] [31, 180] [181,365]	Interim	2.80%	2.45%	2.10%	2.00%
				Final	1.05%	0.85%	0.95%	1.05%
30.30%	22.3%	λ2=0.3*λ1 λ3=0.1*λ1	[0, 30] [31, 180] [181,365]	Interim	2.45%	2.30%	2.25%	2.25%
				Final	0.70%	0.75%	0.65%	0.50%
30.30%	22.3%	λ2=0.2*λ1 λ3=0.2*λ1	[0, 30] [31, 180] [181,365]	Interim	2.85%	2.65%	2.45%	2.30%
				Final	0.85%	0.75%	0.85%	0.80%

*Time intervals are used for both the sample data and the models for the simulations.

The sensitivity analyses are performed to evaluate the robustness of the prespecified piecewise exponential models. The sample data constructed from piecewise exponential distribution based on different time intervals (specified 3.1.9.1) are used to fit the

prespecified piecewise exponential with the prespecified time intervals [0,30], [31,180], [181,365]. The results of the sensitivity analysis for type I error by time interval are shown in Table 6. The results show the type I errors are similar and consistent to the results in Table 5 and reasonably controlled around 2.5%. The simulation results for the type I errors show that the prespecified piecewise exponential models are robust even when constant hazard distributions for the sample data deviate from the prespecified models.

Table 6. Sensitivity Analysis for Type I Error by Time Interval

Simulation Parameters				**Chance of the Posterior Probability for Test Rate minus Control Rate <8% is greater than ξ**				
Test Rate	**Control Rate**	**Hazard Rate**	**Time Interval***	**Time**	**ξ=97.1%**	**ξ=97.3%**	**ξ=97.5%**	**ξ=97.7%**
30.30%	22.3%	λ2=0.4*λ1 λ3=0.2*λ1	[0, 7] [8, 30] [31,365]	Interim	2.95%	2.75%	2.45%	2.15%
				Final	1.20%	0.95%	1.00%	1.10%
30.30%	22.3%	λ2=0.3*λ1 λ3=0.1*λ1	[0, 7] [8, 30] [31,365]	Interim	2.20%	2.15%	2.00%	1.85%
				Final	1.40%	1.15%	1.00%	1.00%
30.30%	22.3%	λ2=0.4*λ1 λ3=0.2*λ1	[0, 7] [8, 90] [91,365]	Interim	3.00%	2.75%	2.60%	2.40%
				Final	1.35%	1.30%	1.25%	1.20%
30.30%	22.3%	λ2=0.3*λ1 λ3=0.1*λ1	[0, 7] [8, 90] [91,365]	Interim	3.10%	3.00%	2.70%	2.40%
				Final	0.90%	0.80%	0.70%	0.60%
30.30%	22.3%	λ2=0.4*λ1 λ3=0.2*λ1	[0, 90] [91, 180] [181,365]	Interim	2.75%	2.70%	2.50%	2.35%
				Final	0.95%	0.85%	0.90%	0.85%
30.30%	22.3%	λ2=0.3*λ1 λ3=0.1*λ1	[0, 90] [91, 180] [181,365]	Interim	2.30%	2.15%	2.00%	1.80%
				Final	0.90%	0.80%	0.80%	0.70%
30.30%	22.3%	λ2=0.4*λ1 λ3=0.2*λ1	[0, 180] [181, 270] [181,365]	Interim	3.15%	2.85%	2.65%	2.35%
				Final	0.75%	0.80%	0.85%	0.80%
30.30%	22.3%	λ2=0.3*λ1 λ3=0.1*λ1	[0, 180] [181, 270] [271,365]	Interim	2.40%	2.30%	2.15%	2.10%
				Final	0.95%	0.90%	0.90%	0.90%

* Intervals used for creating sample data. The models used the intervals: [0,30], [31, 180], [181, 365] as prespecified for the primary endpoint analysis.

The study power (1-minus type II error) is more than 90% at the interim or the final analysis of the primary endpoint from the simulation results as shown in Table 6. The study is sufficiently powered for the primary endpoint analysis.

Table 6. Simulation Results for Study Power

Simulation Parameters				**Chance of the Posterior Probability for Test Rate minus Control Rate <8% is greater than ξ**				
Test Rate	**Control Rate**	**Hazard Rate**	**Time Interval**	**Time**	**ξ=97.1%**	**ξ=97.3%**	**ξ=97.5%**	**ξ=97.7%**
22.30%	22.3%	λ2=0.3*λ1 λ3=0.2*λ1	[0, 30] [31, 180] [181,365]	Interim	95.65%	95.45%	95.15%	94.70%
				Final	97.50%	97.40%	97.05%	96.85%

3.2 ACURATE *Prime* XL Nested Registry Statistical Assessment

The statistical assessment for the ACURATE *Prime* XL Nested Registry is summarized in the sections below.

3.2.1 Statistical Hypothesis – ACURATE *Prime* XL Nested Registry

The statistical hypothesis is that the mean aortic valve pressure gradient at 30 days post implant procedure is less than a performance goal (PG):

H_0: $Gradient_{30D} \geq PG$

H_1: $Gradient_{30D} < PG$

where $Gradient_{30D}$ is the 30-day mean aortic valve pressure gradient for the ACURATE *Prime* XL valve and PG is 15 mmHg.

A one-sample *t*-test will be used to test the one-sided hypothesis at a significance level of 2.5%.

3.2.2 Sample Size Parameters for the ACURATE *Prime* XL Nested Registry

The sample size calculation is based on the following assumptions.

- Expected 30-day mean pressure gradient from ACURATE *Prime* XL = 10 mmHg
- Expected standard deviation = 7 mmHg
- PG = 15 mmHg
- Test significance level (α) = 0.025 (1-sided)
- Power > 90%
- Evaluable number of subjects = Minimum of 40 subjects
- Expected rate of attrition = 20% (8 subjects)
- Planned enrollment of 50 subjects
- The analysis population for the hypothesis testing will be the subject population implanted with the ACURATE *Prime* XL valve.

Note: The expected mean gradient is based on Boston Scientific data on file and published data for large-annulus CoreValve devices (referenced in the protocol synopsis).

3.2.3 Success Criteria for the ACURATE *Prime* XL Nested Registry

If the *P* value from the one-sample *t*-test is < 0.025, the ACURATE *Prime* XL valve will be concluded to have a 30-day mean aortic valve pressure gradient < 15 mmHg. This corresponds to the one-sided upper 2.5% confidence bound of the observed 30-day mean aortic valve pressure gradient being < 15 mmHg.

3.3 Quality of Life (QoL) Endpoint

The quality-of-life endpoint is defined as the percentage of subjects who had moderate or greater improvement of Kansas City Cardiomyopathy Questionnaire Overall Summary Score (KCCQ-OSS) from baseline to 1-year follow-up for subjects who are implanted with the ACURATE valve.

Moderate or greater improvement is defined as more than 10 points improvement of KCCQ-OSS, where the improvement is calculated as KCCQ-OSS at 1-year minus KCCQ-OSS at baseline.

Note: Subjects who died within 1 year and/or had KCCQ-OSS at both baseline and 1-year follow-up will be included in the QoL Endpoint analysis. If a subject died without KCCQ-OSS at 1-year, this subject is considered as not having a moderate or greater improvement of KCCQ-OSS.

3.3.1 Statistical Hypothesis

Percentage of subjects with moderate or greater improvement of KCCQ-OSS from baseline to 1-year follow-up meets the performance goal (PG):

H_0: $P_{moderate\_improve} \leq PG$

H_1: $P_{moderate\_improve} > PG$

where $P_{moderate\_improve}$ is the percentage of subjects with moderate or greater improvement of KCCQ-OSS from baseline to 1-year follow-up. PG is set to be 50%.

3.3.2 Statistical Test Method

A one-sample z-test will be used to test the one-sided hypothesis at a significance level of 2.5%.

3.3.3 Power and Sample Size

The sample size calculation is based on the following assumptions.

- Performance Goal (PG) = 50%
- Expected Rate = 55%
- Test significance level (a) = 0.025 (1-sided)
- Number of subjects to be enrolled > 980
- Expected attrition rate = 20% (196 subjects)

- Number of evaluable subjects > 784
- Power > 80%

Note: Analysis population includes subjects who are implanted with ACURATE valve in the ACURATE IDE Main Randomized Cohort (750 subjects), ACURATE *Prime* XL Nested Registry (50 subjects), Continued Access Study (maximum of 1,000 subjects) and ACURATE Extended Durability Study (approximately 50 subjects).

Note: The moderate or greater improvement was noted in 55% of the subjects with the 95% CI [52.0%, 58.2%] from the meta-analysis random effect model using data from the trials/cohorts listed below. Considering margin of errors and variabilities of the observed data across different studies and operative risks, a 5% margin is applied to the historical rate of 55% to derive the PG. Thus the 50% PG (55% minus 5%) is proposed for a clinical meaningful moderate improvement of the KCCQ-OSS for this study.
Note: Trials/cohorts data included in the meta-analysis:
Boston Scientific REPRISE III study - High/Extreme (CoreValve), CoreValve US Pivotal Extreme Risk Trial (TAVR)[1], PARTNER Trial High Risk (TAVR)[2], PARTNER 2 RCT (TF) - Intermediate (TAVR)[3], PARTNER 3 Low Risk (TAVR)[4].

References:
[1] Suzanne V. Arnold , MD, MHA, et al., Five-Year Clinical and Quality of Life Outcomes From the CoreValve US Pivotal Extreme Risk Trial, Circ Cardiovasc Interv. 2021;14:e010258. DOI: 10.1161/CIRCINTERVENTIONS.120.010258.
[2] Matthew R. Reynolds MD, MSc, et al., Health-Related Quality of Life After Transcatheter or Surgical Aortic Valve Replacement in High-Risk Patients With Severe Aortic Stenosis: Results From the PARTNER, Journal of the American College of Cardiology, Volume 61, Issue 1, 8 January 2013, Pages 108.
[3] Suzanne J. Baron, MD, et al., Health Status Benefits of Transcatheter vs Surgical Aortic Valve Replacement in Patients With Severe Aortic Stenosis at Intermediate Surgical Risk, JAMA Cardiol. 2017 Aug; 2(8): 837–845.
[4] Suzanne J. Baron, MD, MSC, et al., Health Status After Transcatheter Versus Surgical Aortic Valve Replacement in Low-Risk Patients With Aortic Stenosis, JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY, VOL. 74, NO. 23, 2019.

3.3.4 Success Criteria

If the *P* value from the one-sample z-test is < 0.025, it concludes that the QoL Endpoint for the study meets the pre-specified PG of 50%. This corresponds to the one-sided lower 2.5% confidence bound of the observed percentage of subjects with moderate or greater improvement of KCCQ-OSS from baseline to 1-year follow-up is > 50%.

3.3.5 QoL Analysis

The QoL endpoint analysis set will be based on subjects who are implanted with the ACURATE valve from ACURATE valve in the ACURATE IDE Main Randomized Cohort, ACURATE *Prime* XL Nested Registry, Continued Access Study, and ACURATE Extended Durability Study.

The analysis set will include subjects who died prior to the 1-year visit (within 365 days post index procedure) or have both baseline and 1-year KCCQ-OSS.

Note: If a subject died within 365 days post index procedure without KCCQ-OSS at 1-year, this subject is considered as not having a moderate or greater improvement of KCCQ-OSS.
This planned analysis will be performed after the subjects completed 1-year follow-up. Other QoL/KCCQ data may be analyzed using the same analysis set as appropriate.

3.4 Other Endpoints

Additional analyses at interim for other CEC endpoints will be performed on all enrolled Main Randomized Cohort subjects (N=1500) using the piecewise exponential prediction models specified in Section 3.1.5. Event rates and credible intervals will be provided.

For the Main Randomized Cohort (including 4D-CT Cohort subjects), the 4D-CT Cohort, and the Extended Durability Study, the analysis will be conducted separately. Statistical comparisons for these endpoints will be made between the test and the control arms within each cohort. 4D-CT analysis will include the data for subjects who are eligible for the Main Randomized Cohort analysis at the time of the interim analysis.
Analyses for other endpoints will use the standard frequentist methods (details in Section 4.1), such as Chi-square or Fisher's exact for binary variables, T-test for continuous variables, or Log-rank test for time-to-event or Kaplan–Meier survival analysis when they are appropriate, at both the interim and the final analyses. Additional analysis methods may be used when they are appropriated (e.g. Breslow-day test for interaction, ANOVA test or Kruskal-Wallis test for continuous variables, one-sample T-test or McNemar's test for paired analysis).

For the roll-in, the ACURATE *Prime* XL Nested Registry, and the Continued Access Study cohorts, the analyses will be conducted separately, and descriptive analyses will be provided.
Reference:
FDA Guidance for Industry and FDA Staff: Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials, February 2010

4 GENERAL STATISTICAL METHODS

4.1 General Methods

Descriptive statistics and statistical comparisons will be presented on the trial results by treatment group for the Main Randomized Cohort subjects (including 4D-CT Cohort subjects), 4D-CT Cohort, and the Extended Durability Study. Descriptive statistics will be presented separately for roll-in subjects, the ACURATE *Prime* XL Nested Registry subjects, and the Continued Access Study subjects. For continuous variables, summaries will include the sample size (N), mean, standard deviation, minimum, and maximum. Frequency tables will be used to summarize discrete variables. The difference between

comparison groups and its 95% confidence intervals will be calculated. Treatment groups will be compared for randomized subjects using the chi-square test or Fisher's exact test for binary variables and Student's t-test for continuous variables. The appropriate t-test (Pooled or Scatterthwaite methods) will be used based on the test of equality of variances (Levene's test or F-test) between the treatment groups. The chi-square test is used by default; the Fisher Exact test is used in place of the chi-square test when one or both of the following occur: total number of samples ≤40 and/or at least one cell count in the 2X2 table has expected value less than 5. Additional tests such as: log-rank test for Kaplan-Meier rate or chi-square test for 2xn table may be used when they are appropriate. Additional analysis methods may be used when they are appropriated (e.g. Breslow-day test for interaction, ANOVA test or Kruskal-Wallis test for continuous variables, one-sample T-test or McNemar's test for paired analysis).

4.2 Analysis Sets

The primary endpoint and additional measurements will be analyzed on an ITT and an implanted basis.

Among the randomized cohorts (the Main Randomized, the 4D-CT, and the Extended Durability Cohorts), for ITT analyses, all subjects who sign the IRB/REB/HREC/ IEC -approved study ICF, are enrolled in the trial, and are randomized will be included in the analysis, whether or not an assigned study valve (ACURATE or Control) was implanted. For Implanted analyses, ITT subjects who had the assigned (Test versus Control), randomized study valve implanted will be included in the analysis. For the implanted randomized cohort analysis set, if a subject receives 2 different valve types from 2 different manufacturers, the subject will be excluded from the implanted analysis.

With the roll-in cohort, the ACURATE *Prime* XL Nested Registry cohort, and the Continued Access Study, the subject is considered enrolled in the trial when there is an attempt made to insert the ACURATE Transfemoral Aortic Valve System into the subject's femoral artery. For ITT analyses, all subjects who sign the IRB/REB/HREC/ IEC -approved study ICF and are enrolled in the trial will be included in the analysis sample, regardless of whether the study device was implanted. The implanted population includes all subjects who sign an ICF and are implanted with an ACURATE valve.

For the randomized cohorts (the Main RCT, the 4D-CT, and the Extended Durability cohorts), the subjects will be analyzed for the ITT and implanted analysis sets. The primary analysis for the primary endpoint will be based on the ITT analysis set.

For ITT analysis set, events starting from the randomization date will be included in the analysis for the randomized cohorts (the Main RCT, the 4D-CT, and the Extended Durability cohorts) and events starting from the procedure date will be included in the analysis for the roll-in cohort, the ACURATE *Prime* XL Nested Registry cohort, and the Continued Access Study. For the Implanted analysis sets, events starting from the procedure date will be included in the analysis.

With the 4D CT cohort, the 4D CT imaging analysis is based on the implanted analysis set. 4D CT data will be pooled with available ACURATE neo2 PMCF data. The pooled data and individual 4D CT cohort will be analyzed.

The analysis set at the formal interim is the subset of randomized subjects who had completed 30-day, 6-month, and/or 1-year follow-up as defined in Section 3.1.5.1. The analysis set (N=1500) at the final will include all enrolled subjects.

4.3 Control of Systematic Error/Bias

All subjects who have met the inclusion/exclusion criteria, received a positive recommendation from the CRC, and signed the Informed Consent Form will be eligible for enrollment in the study. The center heart team's assessment of TTE measurements before device placement will contribute to the determination of subject eligibility for the study.
To control for inter-observer variability, data from independent core laboratories will be used for analysis. These include an echocardiography core laboratory and a CT/angiography core laboratory to assess all data using standard techniques. Clinical endpoints will also be adjudicated by an independent CEC.
The study team is blinded to aggregated data reports prior to the official study unblinding at the formal interim analysis.

4.4 Randomization Schedule

For the Main Randomized Cohort and Extended Durability Study, a computer-generated list of random treatment allocations for each cohort (i.e., a randomization schedule) will be used to assign subjects in the ACURATE IDE trial to treatment in a 1:1 ratio of ACURATE to Control. Randomization will be stratified by center and by intended control device (SAPIEN or CoreValve). Additional information is provided in the study Manual of Operations.

4.5 Number of Subjects per Investigative Site

Enrollments shall not exceed 18% of total enrolled RCT subjects at any individual investigative site for the ACURATE IDE study. There is no set minimum number of subjects to be enrolled per site. "Small sites" will be -removed for poolability analyses (see Section 5.5.1).

5 ADDITIONAL DATA ANALYSES

5.1 Adjudicated CEC Event Analysis

For clinical events, the cut-off for events for 30-day endpoints will be 30 days, for 6-month endpoints it will be 180 days, for 1-year endpoints it will be 365 days, and for endpoints at 2–10 years it will be 365 days times the number of years. For events at discharge or 7 days post-procedure, the cut-off for events will be the earlier date of discharge or 7 days post-procedure for each subject. The number of evaluable subjects for the CEC event includes subjects who have CEC confirmed events or are in study at least 23, 150, or 335 days (defined as sufficient follow-up in Section 7.3.3) for 30-day, 6-

month, or 1-year outcome analysis, respectively. For analysis at 2-10 years, the sufficient follow-up is defined as “number of years times 365 minus 45”. The number of evaluable subjects for the CEC event at discharge will include all subjects in the specific analysis set.

The Kaplan-Meier rates for all CEC Events will be provided in analyses at 30 days, 6-month, and 1-10 years. The survival analysis (including accumulated event curve) for time-to-event endpoint will be provided for selected CEC events, and log-rank test p-value and wilcoxon test p-value will be provided for comparing two survival curves.

Event to be analysed include the following:

- Mortality: all-cause, cardiovascular, and non-cardiovascular
- Stroke: disabling and non-disabling
- Myocardial infarction (MI): periprocedural (≤72 hours post index procedure) and spontaneous (>72 hours post index procedure)
- Bleeding: life-threatening (or disabling) and major (through 5 years)
- Acute kidney injury (AKI; ≤7 days post index procedure): based on the AKIN System Stage 3 (including renal replacement therapy) or Stage 2
- Major vascular complications (through 5 years)
- Repeat procedure for valve-related dysfunction (surgical or interventional therapy)
- Hospitalization for valve-related symptoms or worsening congestive heart failure (NYHA class III or IV)
- New permanent pacemaker implantation resulting from new or worsened conduction disturbances
- New onset of atrial fibrillation or atrial flutter
- Coronary obstruction: periprocedural (≤72 hours post index procedure)
- Ventricular septal perforation: periprocedural (≤72 hours post index procedure)
- Mitral apparatus damage: periprocedural (≤72 hours post index procedure)
- Cardiac tamponade: periprocedural (≤72 hours post index procedure)
- Valve migration
- Valve embolization
- Ectopic valve deployment
- Transcatheter aortic valve (TAV)-in-TAV deployment
- Prosthetic aortic valve thrombosis
- Prosthetic aortic valve endocarditis

5.2 Other Clinical Assessments

The following assessment will be evaluated.

- Device performance endpoints peri- and post-procedure:

  - Successful vascular access, delivery and deployment of the study valve and successful retrieval of the delivery system
  - Grade of aortic valve regurgitation: paravalvular, central and combined
- Device success, defined as absence of procedural mortality, correct positioning of a single transcatheter valve in the proper anatomical location, and intended performance of the study device (indexed effective orifice area [iEOA] >0.85 cm^2/m^2 for BMI <30 kg/cm^2 and iEOA >0.70 cm^2/m^2 for BMI ≥30 kg/cm^2 plus either a mean aortic valve gradient <20 mm Hg or a peak velocity < 3m/sec, and no moderate or severe prosthetic valve aortic regurgitation)
- Additional indications of prosthetic aortic valve performance as measured by transthoracic echocardiography and assessed by an independent core laboratory, including EOA, mean and peak aortic gradients, peak aortic velocity, and grade of aortic regurgitation
- Functional status as evaluated by New York Heart Association (NYHA) classification
- Neurological status as determined by the following:
  - National Institutes of Health Stroke Scale (NIHSS) conducted by a neurology professional or certified personnel at discharge and 1 year
  - Modified Rankin Scale (mRS) conducted by a neurology professional or certified personnel at discharge and all follow-up visits
  - Neurological physical exam conducted by a neurologist, neurology fellow, neurology physician assistant, or neurology nurse practitioner in all subjects where stroke is suspected

Health status as evaluated by Kansas City Cardiomyopathy and SF-12 Quality of Life (QOL) questionnaires at baseline, 1 month, 1 year, and 5 years.

For subjects in the CT Imaging Substudy, assessments using 4D CT at 30 days and 1 year will be done as listed below. Data will be evaluated by an independent CT core laboratory.

- Assessment of leaflet mobility
- Assessment of hypoattenuated leaflet thickening (HALT)
- Assessment of leaflet thrombosis

5.3 Interim Analyses

Two interim analyses are planned. The administrative interim analysis will be conducted when the first 350 Main Randomized Cohort subjects have completed 1-year follow-up visits. The formal interim analysis will be carried out after enrollment in the Main Randomized Cohort is completed. This formal interim analysis will be conducted on the full N=1500 subjects in the Main Randomized Cohort after the first 1050 subjects in the Main Randomized Cohort have completed 1-year follow-up. The piecewise exponential model based on outcomes among these subjects will be used to estimate the 1-year results

by treatment group for the remaining enrolled subjects. A final analysis on hypothesis of the primary endpoint will be performed on all Main Randomized Cohort subjects with completed 1-year data if non-inferiority cannot be claimed at the formal interim analysis

A Data Monitoring Committee (DMC) will monitor safety events. The DMC reports will be generated according to the DMC charter.

The 4D-CT Cohort (pooling with PMCF 4D-CT subjects) data will be analyzed for the subjects who have completed 30 days, 6 months, and 1-year follow-ups at the time of the formal interim analysis, as appropriate.

For roll-in cohort, data will be analyzed for the subjects who have completed 1 year follow-up and additional data may be provided at the time of the formal interim analysis, as appropriate.

For the ACURATE *Prime* XL Nested Registry cohort, data will be analyzed when all enrolled subjects have completed 30 days follow-up. Additional data may be provided at the time of the formal interim analysis, as appropriate.

Analyses of the Extended Durability Study and the Continued Access Study are not expected to occur at the time when the formal interim analysis is performed on the Main Randomized Cohort. Time of the analyses for the two sub-studies will be based on the data availability.

5.4 Subgroup Analyses for the Main Randomized Cohort Subjects

Primary and pre-specified additional endpoints will be summarized for the following subgroups of the Main Randomized Cohort subjects. The subgroup analyses will be performed in the ITT analysis population.

- Age (age <75 years and >=75 years)
- Race and ethnicity (categories with more than 100 enrolled subjects)
- Gender (male and female)
- Valve type (ACURATE neo2, SAPIEN, CoreValve)
- Subject operative risk (extreme/high, intermediate, low)
- Valve size (Small, Medium, Large)

Treatment groups will also be compared in the age, race, and gender subgroup analysis. No adjustments for multiple comparisons will be made. Additional analyses may be performed as appropriate.

In the 4D CT Imaging Substudy, computed tomography data will be analyzed after 200 subjects have reached 1-year follow-up.

5.5 Justification of Pooling

5.5.1 Pooling Study Centers

An assessment of poolability across sites for the Main Randomized Cohort will be performed. In the analysis, centers with fewer than 5 subjects enrolled in either ACURATE arm or Control arms in the study will not be included in the site poolability

analysis using the logistic regression model. Descriptive statistics for the primary endpoint for each individual center will be presented in the report.

Main effects for the study centers and treatment and the interaction of the center by treatment will be included in a logistic regression model with the primary endpoint as the outcome. If the P value for center by treatment interaction is > 0.15, it can be concluded that the treatment effect is not different across the study centers and the data can be pooled across study centers. If the resulting P value is ≤ 0.15, the interaction will be examined if it is a quantitative or qualitative interaction, and further exploratory analysis will be performed to identify potential outlier study centers or potential covariates that may explain treatment differences between the centers. The cases from the individual centers will be reviewed and the poolability analysis may be performed by removing the individual centers or by adjusting these covariates. Under this circumstance, descriptive statistics of covariates and the primary endpoint outcomes by treatment group for each site will be provided.

5.6 Multivariable Analyses

Univariate and multivariate analyses will be performed to assess the effect of potential predictors on the primary endpoint and additional endpoints. The following variables will be analyzed as possible predictors:

Category	**Covariates**
Treatment	Treatment group (ACURATE=1 vs Control=0)
Cohort	4D-CT Cohort
Demographics and Baseline Characteristics	Age at time of consent, Gender, Race, BMI, NYHA, STS Score, EuroSCORE II, Operative risk
General Medical History	Medically treated diabetes mellitus, History of hyperlipidemia, History of hypertension, History of peripheral vascular disease, Severe Liver Disease/Cirrhosis, History of Dialysis Dependent Renal Failure, History of Chronic Obstructive Pulmonary Disease, Severe lung disease, Severe pulmonary hypertension
Cardiac History	History of Coronary Artery Disease, History of Myocardial Infarction, History of congestive heart failure, Prior Balloon Aortic Valvuloplasty, Current Anginal Status, History of Percutaneous Coronary Intervention (PCI), History of Coronary Artery bypass Graft (CABG) Surgery, Hostile Chest, Porcelain Aorta, Prior chest radiation therapy, Neuromuscular disease that creates risk for mechanical ventilation or rehabilitation after surgical aortic valve replacement, Orthopedic disease that creates risk for rehabilitation after surgical aortic valve replacement, History of Atrial Fibrillation, History of Atrial Flutter, Prior Pacemaker Implant
Neurological History	History of cerebrovascular accident (TIA/Stroke), Right Carotid Artery Stenosis (>= 80%), Left Carotid Artery Stenosis (>= 80%), Prior Carotid Endarterectomy / Carotid Artery Stenting

Frailty	Time to walk 5 meters, Use of wheelchair, Number of falls in the past 6 months, Maximum Grip Strength Average, Katz Basic Activities of Daily Living Score
Medications	DAPT at baseline (excluding loading dose), Anticoagulant
Echo Aortic Factors (Core lab)	Aortic valve area AVA (VTI), Mean Aortic valve gradient, Aortic valve peak velocity, LVEF, Transvalvular aortic regurgitation, Doppler Velocity Index (DVI) (VTI), Systolic pulmonary pressure (sPAP)
CTA (Core Lab)	Annular area, LVOT area
Peri-procedural Characteristics	Post-dilation balloon performed
Procedural Characteristics	Total procedure time, Anesthesia type, BAV used during index procedure, Pre-dilation maximum balloon diameter, Total contrast media used for procedure, Total fluoroscopy time, TEE used during the implant procedure, Embolic protection device

Note: Covariate may not be included in multivariable model due to convergence criterion (GCONV=1E-8) is not satisfied. Covariate may not be included in multivariable model due to missing data on more than 10% of subjects.

Possible predictors will be identified using univariate models with baseline or other identified covariates. These identified predictors from the univariate models with $p \leq 0.20$ will also be modeled multivariately using a stepwise procedure in a logistic regression model. The significance level thresholds for entry and exit of independent variables into the multivariate model will be set at 0.1.

From the final models, predictors will be listed in ascending order of p-value. Univariate analyses will be performed overall as well as separately for each treatment group for the Main Randomized Cohort subjects.

5.7 Other Analyses

The general statistical methods (see Section 4.1) will be applied appropriately for the following analyses. The ITT and implanted analysis sets (see Section 4.2) will be used for following analyses.

5.7.1 Baseline Characteristics/Medical History

Baseline data will be summarized to assess subject demographics, clinical history, risk factors, and pre-procedure characteristics.

5.7.2 Post-Procedure Endpoints

Post-procedure information will be collected at regularly scheduled follow-up examinations as detailed in the clinical trial schedule in the protocol.

5.7.3 Subject Disposition and Device Disposition

Subject disposition and Device Disposition (e.g., number completing the study, number lost-to-follow-up) will be summarized with frequency tables.

5.7.4 Medication

Antiplatelet and anticoagulant medications will be analyzed at baseline and follow-up visits.

5.7.5 Site reported AE/SAE/UADE

Site reported adverse events will be summarized with frequency tables.

5.7.6 Protocol Deviations

Summary analysis will be provided by protocol deviation categories.

5.7.7 Other Clinical Assessment at various points

- NYHA classification
- 12-lead electrocardiogram (ECG)
- Risk assessments: Society of Thoracic Surgeons (STS) score, euroSCORE II
- Frailty, disability, and comorbidity assessments
- A CT angiogram
- NIH Stroke Scale (NIHSS)
- Modified Rankin Scale (mRS) score
- Quality of Life (QOL)
- SF-12 QOL Questionnaires
- 4D CT at 30 days and 1 year

5.7.8 COVID-19 related analyses

Data collected related to COVID-19 will be assessed. If there is significant impact of COVID-19 on results, additional analyses (eg: sensitivity) will be performed to assess the impact of COVID-19 on the study results according to regulatory guidance.

5.8 Changes to Planned Analyses

Any changes to the planned statistical analyses made prior to performing the analyses will be documented in an amended SAP approved before performing the analyses. Changes from the planned statistical methods after performing the analyses will be documented in the clinical study report along with a reason for the deviation.

6 VALIDATION

All clinical data reports generated per this plan will follow the Global WI: Clinical Data Reporting Validation (PDM 90702587).

7 PROGRAMMING CONSIDERATIONS

7.1 Statistical Software

Statistical data review will be performed by the sponsor. Statistical analyses will be performed using SAS System software, version 9.2 or later (Copyright © 2000 SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513, USA. All rights reserved).

7.2 Format of Output

Results of analysis will be output programmatically to Word documents from SAS with no manual intervention. All output for the final statistical report will be in the form of a Word document containing tables, figures, graphs, and listings, as appropriate.

7.3 Rules and Definitions for Calculated Variables

7.3.1 Transthoracic Echocardiographic (TTE) Variables

One transthoracic echocardiographic study will be performed for each visit. If multiple transthoracic echocardiographic studies are performed for the same visit, the latest study performed for each visit will be used for analysis.

7.3.1.1 Body Mass Index (BMI)

Analysis approach: Body Mass Index is calculated for each visit.

Weight (Kg) = Weight (lbs) / 2.20462262.
Height (cm) = Height (in) / 0.393700787

$$BMI = \frac{Weight\,(Kg) \times 10000}{\left(Height\,(cm)\right)^2}$$

7.3.1.2 Body Surface Area (BSA)

Analysis approach: Body Surface Area (BSA) is calculated using the following formula for each visit if data is available.:

$$BSA\,(m^2) = \sqrt{(Height\,(cm) x Weight(Kg))/3600}$$

7.3.1.3 Indexed Aortic Valve Area (iAVA) or Indexed Effective Orifice Area (iEOA).

Effective Orifice Area (EOA) is synonymous with Aortic Valve Area (AVA).

Analysis approach:

Indexed Aortic Valve Area (iAVA) or Indexed Effective Orifice Area (iEOA) is calculated for each visit.

iAVA(cm2/m2) = iEOA(cm^2/m2) = AVA (TVI) (cm^2) / BSA (m^2),

where AVA (TVI) is the aortic valve area for a specific visit and BSA is the body surface area for the same specific visit under analysis.

7.3.2 Days to Last Follow-up

Valid Data Sources

- Adverse Event Form
- Hospitalization Form
- Procedure Form
- Date of Visit Form
- CEC data
- End of Study

Valid Data Points

- Adverse event dates are “Onset date”, “If Hospitalization, is this a new hospitalization, If Yes, admission date”, and “If Yes, discharge date” from the Adverse Event Form.
- Admission and Discharge dates are “Admission date” and “Discharge date” from the Hospitalization Form.
- Procedure date is “Date of Procedure” from the Procedure Form.
- Enrollment date is “Date of Enrollment” from the Screening and Additional Informed Consent Form
- Follow-up visit date is “Date of Visit” from the Date of Visit Form at each of the visits (discharge or 7 days post-procedure, 30 days, 6 months, and 1 to 10 years post-procedure).
- CEC event date – date of event as adjudicated by the CEC
- End of Study – dates of subject withdrawal consent and death.

Last follow-up date will be the latest of the following dates for each subject:

adverse event onset date,
admission and discharge dates from hospitalization,
procedure date,
enrollment date
discharge or follow-up visit date, and
CEC event date.

If a subject died or withdrawn consent, the date of death or withdrawal (which even came first) should be the last follow-up date.
Follow-up days will be calculated for intent-to-treat analysis set and implanted analysis set

Day 0 is randomization date (ITT) or the procedure date (Implanted)

Days to last follow-up = last follow-up date – randomization date (ITT)/procedure date (Implanted).

Days to (event or last known status) = (event or status) date - randomization date (ITT)/procedure date (Implanted).

7.3.3 Event Rates

The calculation of binary rates will be the same for any endpoint and time point in regard to the appropriate numbers of days as indicated below in Table 1

Table 1. Days Post-procedure to Event and for Adequate Follow-up for ITT and Implanted Analysis Sets.

Follow-up Visit	Maximum Days to Event from Randomization/Procedure*	Days for Adequate Follow-up from Randomization/Procedure**
30 Days	30	23
6 Months	180	150
12 Months	365	335
2 Years	730	685
3 Years	1095	1050
4 Years	1460	1415
5 Years	1825	1780
6 Years	2190	2130
7 Years	2555	2495
8 Years	2920	2860
9 Years	3285	3225
10 Years	3650	3590

* Target date for the follow-up visit.
** Start of the follow-up visit window.
Note: For ITT analysis, the Days are from randomization date. For Implanted, the Days are from procedure date (for non-RCT cohorts, the Days are from procedure date).

Binary event rates (proportions) are calculated on a per subject basis.

Event rates through a follow-up visit are calculated using the following for inclusion in the denominator and numerator:

- Denominator:

  Subjects in the specific analysis set count in the denominator with one of the following:

  - Subject experiences any CEC adjudicated event ≤ maximum number of days as specified in Table 1, as appropriate or
  - date of last follow-up ≥ days for adequate follow-up post-procedure from Table 1, as appropriate:

- Numerator:

  Subjects in the specific analysis set count in the numerator if the subject experiences specified event ≤ maximum number of days as specified in Table 1, as appropriate.

7.3.4 Missing Dates

When calculating rates of adverse events, missing and partial onset dates will be handled as shown in the table below.

Partial Date	**Action Taken**
Entire adverse event onset date is missing	The procedure date will be used for the onset date.
The month and the day of the month are missing but the year is available	January 1st will be used for the month and day of the onset date. However, if the imputed date falls before the procedure date, then the procedure date will be used for the onset date.
Day is missing, but the month and year are available	The 1st will be used as the day of the onset date. However, if the imputed date falls before the procedure date, then the procedure date will be used for the onset date.

For determination of medication history following imputation will be applied for the missing and partial start or stop date of medications.

Medication start or stop date (Partial)	**Imputed Date***
Day is missing, but the month and year are available	Consider 1st Day of the month.
Both Day and Month are missing	Consider 1st Day of the month for Day and arbitrarily assign July for the missing month
Year is missing	No imputation

*If imputed date falls after patient last visit date, then the patient last visit date will be used for the medication last date. The medication start date can be earlier than procedure/randomization dates.

Note: Imputed dates will not be used for “last date” calculation

8 REVISION HISTORY

Document Revision Number	Template Number and Version	Section	Change	Reason for Change
A	90702621 Rev/Ver AE		New SAP	N/A
B	90702621 Rev/Ver AE	1, 3, 4, 5	Revisions to IC#4, primary endpoint analysis datasets, primary endpoint backup option, administrative analysis.	Revisions to IC#4 made to align with Protocol amendment vC. Additional text added and modified throughout per FDA request to clarify that the primary analysis for FDA regulatory submission will be based on the "patient-level" anonymized subject dataset. Typo corrected in the primary endpoint backup option to clarify "disabling stroke" should be "all stroke". Clarified that for the administrative analysis, subgroup for race and ethnicity data is not available in SCOPE I/II study and that the subgroup analysis will be done for ACURATE IDE only.
C	90702621 Rev/Ver AE	3, 3.2.4	Formatting changes	Administrative font and formatting changes; no changes to language or text.
D	90702621 Rev/Ver AE	1	**Roll-In Cohort:** A non-randomized roll-in phase… will perform at least 2 roll-in cases before commencing treatment in the randomized cohort.	Clarify centers must have performed 2 roll-in cases before treating subjects in the randomized cohort.
		1, 3	Subjects will be enrolled at up to 65 centers… There will be up to 630 subjects in ACURATE IDE	Support faster enrollment
		1	<u>Aspirin</u> A loading dose of aspirin… is recommended for subjects... The loading dose should be administered prior to the implant procedure…. <u>$P2Y_{12}$ Inhibitor</u> A loading dose of a $P2Y_{12}$ inhibitor… is recommended for	Current standard of care; updated text for clarification

			subjects... The loading dose should be administered prior to the implant procedure…. **Note 5:** If a subject requires chronic anticoagulation, either a $P2Y_{12}$ inhibitor or aspirin is recommended prior to and required after the implant procedure in addition to the anticoagulant therapy (but treatment with both aspirin and a $P2Y_{12}$ inhibitor after the implant procedure is not required).	
		1	Subjects at low surgical risk were subsequently approved for TAVR with commercially available devices in the United States.	Added to reflects the updated FDA approval of TAVR in low-risk patients.
		5.5	Valve type (ACURATE *neo*, ACURATE *neo2*, SAPIEN, CoreValve)	Clarify subgroup analysis
E	90702621 Rev/Ver AE	Any section involved	Change “Primary safety endpoint” to “Primary endpoint”	Protocol updates on the Primary endpoint to All death/All stroke/rehospitalization at 1-year
E	90702621 Rev/Ver AE	Any section involved	Remove SCOPE I and SCOPE II related analyses	Protocol updates, SCOPE I and SCOPE II data will not be used in the analyses.
E	90702621 Rev/Ver AE	Any section involved	Sample size calculation change from Intermediate risk to All risk, the sample size parameter and sample size calculation updates	Updated protocol on study design
E	90702621 Rev/Ver AE	Any section involved	Remove the original frequentist analysis method and add the Bayesian methods	Updated protocol on study design change
E	90702621 Rev/Ver AE	Any section involved	Add CT cohort	Updated protocol

E	90702621 Rev/Ver AE	Section 3.1.4.5	Changed interim analysis plan as one administrative analysis and formal interim, and the analysis method to be Bayesian for the formal interim analysis	Updated protocol
E	90702621 Rev/Ver AE	Section 3.1.4.7	Updated sensitivity analyses based on the Bayesian method	Update protocol and statistical method using Bayesian
E	90702621 Rev/Ver AE	Synopsis - Sample Size Parameters for the Primary Endpoint, Section 3.1.2	All design parameters and sample size.	Study design change and SAP updates based current study protocol K
E	90702621 Rev/Ver AE	Section 3.1.5	Provide Bayesian Design Operating Characteristics from simulation	For the study design justification of type I and type II errors
E	90702621 Rev/Ver AE	Section 3.2	Specify Bayesian prediction and analysis for additional CEC event endpoint	Define the additional endpoints that use Bayesian analysis.
E	90702621 Rev/Ver AE	Synopsis – study design, Method of Assigning Subjects to Treatment Section 3, 4.2	Add “4D CT Imaging Substudy” “Data will be summarized separately for the randomized subjects in the 4D CT Imaging Substudy.” “With the 4D CT cohort, the 4D CT imaging analysis is based on the implanted analysis set”	Added 4D CT analysis
E	90702621 Rev/Ver AE	Synopsis - Additional Measurements, Section 5.2	Add “4D CT endpoints”	Study design change and SAP updates based current study protocol

E	90702621 Rev/Ver AE	Synopsis - Planned Number of Subjects and Planned Number of Investigational Centers/ Countries	Sample size change: RCT changed from 500 to 1500 Roll-in changed from 130 to 170 due to number of site increased from 65 to 85	Study design change and SAP updates based current study protocol
E	90702621 Rev/Ver AE	All sections involved, and Follow-up Schedule	Study 5 years follow-up changed to 10 years. And updated the description on how the visits will be performed	SAP updates based current study protocol K
E	90702621 Rev/Ver AE	Synopsis - Additional Exclusion Criteria	Add AEC1, AEC2, and AEC3 that were not included in the SAP version D	SAP updates based current study protocol K
E	90702621 Rev/Ver AE	Synopsis - Analysis Sets, Section 4.2	Remove “For the ITT randomized cohort analysis set, if a subject receives 2 valves, the subject is assigned to the group corresponding to the first valve received”	Correct to the ITT definition in SAP version D.
E	90702621 Rev/Ver AE	Section 3, 4.6, 5.6	Remove analyses related to SCOPE I/II data. Remove Primary Safety Endpoint Remove Backup Options	Study design change.
E	90702621 Rev/Ver AE	Section 3.1	Study design and method change from Standard Frequentist approach to Bayesian approach	Protocol K updates on study design and analysis methods
E	90702621 Rev/Ver AE	Section 4.1	Add “The appropriate t-test (Pooled or Scatterthwaite methods) will be used based on the test of equality of variances between the treatment groups.”	Add specification for t-tests
E	90702621 Rev/Ver AE	Section 5.5	Remove subgroup for ACURATE neo	Not applicable for the IDE study.

E	90702621 Rev/Ver AE	Section 5.3	Update pooling analysis – removing the pooling analysis related to SCOPE I/II studies	Updated protocol on the study design.
E	90702621 Rev/Ver AE	Section 3.1.5	Add details on how the predictive probability will be used to predict outcomes for patients who don't have full follow-up and how exactly the predictive probability will be used for the non-inferiority test. Clarify that The piecewise exponential survival model will be performed separately for each treatment group.	FDA comments
E	90702621 Rev/Ver AE	Section 3.1.7	Clarify that the piecewise exponential survival model will be performed separately for each cohort for the sensitivity analyses. List the cohorts that will fit separate piecewise exponential models. List the details how the predictive probability will be used in the analyses.	FDA comments
E	90702621 Rev/Ver AE	Section 3.1.9.1	Clarify that the piecewise exponential survival models were performed separately for each treatment group in the simulations	FDA comments
E	90702621 Rev/Ver AE	Section 5.5.1	Descriptive statistics of the Primary Endpoint for each individual site, including small sites, be presented in the clinical report, but that the very small sites be excluded from the poolability analysis	FDA comments

E	90702621 Rev/Ver AE	Section 3.1.5	Update the reference: Wilber DJ, et al. *JAMA. 2010;303(4):333-340*	Cite the original paper
E	90702621 Rev/Ver AE	Section 3.1.4	updated the generic formulas to the study specific formulas with details	FDA comments
E	90702621 Rev/Ver AE	Section 3.1.5	Updated based on updated formulas in Section 3.1.4 and add more specifications, Update the approach used to estimate the primary endpoint rates	FDA comments
E	90702621 Rev/Ver AE	Section 3.1.7	Updated formulas based on updated formulas in Section 3.1.4	FDA comments
E	90702621 Rev/Ver AE	Section 3.1.9.2	Updated the simulation results based on the updated approach in Section 3.1.5	FDA comments
E	90702621 Rev/Ver AE	Section 5.7.8	Add potential COVID-19 related analyses	BSC requirement
E	90702621 Rev/Ver AE	Section 3.14, 3.15, 3.16, 3.17	Change method: models (6) by treatment group and time-interval to models (18) by risk group, treatment group, and time interval.	FDA comments
E	90702621 Rev/Ver AE	Section 3.1.5.3	Fixed the error for calculating the primary endpoint rate from hazard rate. The time duration changed from 160 to 150 days	FDA comments
E	90702621 Rev/Ver AE	Section 3.1.9	Change simulation setup for the updated models, update the simulation results. Added the simulation results for sensitivity analysis by time intervals for type I error.	FDA comments

E	90702621 Rev/Ver AE	Section 7.3.4	Add specification for imputing missing dates for AE and Medication (no impact on the primary endpoint analysis)	BSC review comments
E	90702621 Rev/Ver AE	Synopsis - Sample Size Parameters for the Primary Endpoint	Change "the Bayesian method will be used to perform the hypothesis testing from the posterior distributions of the estimated parameters" to "the posterior distributions for the parameters of interest will be used to evaluate the hypothesis testing"	FDA comment – to clarify that posterior distributions of parameters, but not the estimated parameters (which are not random variables)
E	90702621 Rev/Ver AE	Section 3.1.4.3	MCMC will be based on the log likelihood function and lambdas will be estimated by time interval.	Clarify FDA's comments on how the MCMC will be performed.
E	90702621 Rev/Ver AE	Section 3.1.7	Clarify the missing data definition for the tipping point analysis Define the weights used for sensitivity analysis for the control group	FDA comments
F	90702621 Rev/Ver AF	Section 3.1.7	"The time interval sensitivity analyses will be based on the exact same data set that the main primary analysis will be conducted on"	Clarification to FDA question
F	90702621 Rev/Ver AF	Section 3.1.9.1	"Assumption of number of subjects available for each time interval at the time of the interim analysis by risk stratum for each treatment group (total N=750 per treatment group)"	Clarification to FDA question

G	90702621 Rev/Ver AF	All Sections that related to the ACURATE *Prime* XL Nested Registry cohort Section 3.2	Add analysis plan for the the ACURATE *Prime* XL Nested Registry cohort: Section 3.2 Sample size justification and hypothesis test for the cohort	Study design change with added new cohort (Protocol version L)
G	90702621 Rev/Ver AF	Section 5.4	Valve size (Small, Medium, Large)	Protocol version L updates for the additional subgroup.
G	90702621 Rev/Ver AF	Section 3.3	Additional analysis methods may be used when they are appropriated (e.g. Breslow-day test for interaction, ANOVA test or Kruskal-Wallis test for continuous variables, one-sample T-test or McNemar's test for paired analysis).	To specify some additional statistical methods that may be used during analysis when they are considered needed and appropriate.
G	90702621 Rev/Ver AF	Section 4.2	4D CT data will be pooled with available ACURATE neo2 PMCF data. The pooled data and individual 4D CT cohort will be analyzed.	In additional to the 4D CT RCT cohort analysis, BSC plan to pool additional data from PMCF neo2 data for the 4D CT patient analysis.
H	90702621 Rev/Ver AF	Entire SAP ver G (at the places that involves cohorts, statistical methods, and analysis sets)	Add the extended durability (N=100) and continued access study (N=1000) cohorts. Change the "randomized cohort" to the "main randomized cohort" to differentiate it from the randomized "extended durability" cohort.	Protocol ver L to ver M. Specify what analysis on what data from which cohort will be performed.

H	90702621 Rev/Ver AF	Section 3.3	Add “Statistical comparisons for these endpoints will be made between the test and the control arms within each cohort” for the extended durability cohort	Protocol ver L to ver M. for the randomized extended durability cohort.
H	90702621 Rev/Ver AF	Section 4.1	Additional analysis methods may be used when they are appropriated (e.g. Breslow-day test for interaction, ANOVA test or Kruskal-Wallis test for continuous variables, one-sample T-test or McNemar’s test for paired analysis)	specify additional statistical methods for study cohorts.
H	90702621 Rev/Ver AF	Section 4.3	The study team is blinded to any aggregated data reports prior to the official study unblinding at the formal interim analysis	Clarify the RCT blinding requirement.

H	90702621 Rev/Ver AF	Section 5.3	The 4D-CT cohort (pooling with PMCF 4D-CT subjects) data will be analyzed for the subjects who have completed 30 days, 6 months, and 1-year follow-ups at the time of the formal interim analysis, as appropriate. For roll-in cohort, data will be analyzed for the subjects who have completed 1 year follow-up and additional data may be provided at the time of the formal interim analysis, as appropriate, as appropriate. For the ACURATE *Prime* XL Nested Registry cohorts, data will be analyzed when all enrolled subjects have completed 30 days follow-up. Additional data may be provided at the time of the formal interim analysis, as appropriate. Analyses of the extended durability cohort and the continued access study are not expected to occur when the formal interim analysis is performed on the main randomized cohort. Data from these cohorts may not be analyzed at the time of the formal interim analysis due to the data availability.	Specify what analyses for what cohort data will be provided at the time of the formal interim analysis for the main randomized cohort.

H	90702621 Rev/Ver AF	Synopsis	Note: Addition to the criterion of the 1-year completion of the first 1050 subjects in the Main Randomized Cohort, the formal interim analysis will only be conducted after at least 80% low risk subjects in Main Randomized Cohort have completed the 6-month follow-up, whichever comes later.	FDA asked for 80% low risk subjects reached 6-month for the PMA submission.
H	90702621 Rev/Ver AF	Section 3.1.2	after first 1050 subjects in the Main Randomized Cohort have completed 1-year follow-up or at least 80% low risk subjects in the Main Randomized Cohort have completed 6-month follow-up, whichever comes later	FDA asked for sufficient low risk subjects reached 6-month for the PMA submission
H	90702621 Rev/Ver AF	Section 3.1.5.1 Add section 3.1.5.2	1. Enhance the definition for the completion of a follow-up 2. Define the cutoff date and the analysis set for different time intervals for the piecewise exponential models Formal interim analysis cutoff date and analysis sets	Clarify the definition for completion of a follow-up for the case who has not follow-up visits. Pre-specify (before unblinding) the definition for the cutoff date, which will uniquely determine the analysis sets for the piecewise exponential models to avoid selection bias.

H	90702621 Rev/Ver AF	Section 3.1.7	The tipping-point analysis is based on the exact same data set used in the main primary endpoint analysis. The missing cases will be imputed by the chronological order of the imputed days to the primary endpoint event and then by the randomization date or procedure date. For this sensitivity analysis, each time interval (such as: 0-30 days, 31-180 days, and 181-365 days) will be partitioned into three intervals with approximately equal number of events. The Proc PHREG procedure will be used. For each individual model, if the model with eight intervals by default doesn't exist, number of intervals will be reduced to four intervals. If model with four intervals still doesn't exist, the number of intervals for the model will be reduced by one each time until a valid model is identified. Partitions and number of time intervals between different models could be different. Sensitivity analysis may be performed using the observed weights for the operative risk cohorts. The weights are proportional to the numbers of observations between the operative risks (extreme/high,	Add details on tipping-point analysis details on what data will be imputed and by what order, and the sensitivity analyses.

			intermediate, and low) from the observed data that are used to fit each model. The sensitivity analysis is based on the exact same data set used in the main primary endpoint analysis.	
I	90702621 Rev/Ver AF	Section 3.3	Add a new section 3.3 QoL Endpoint	To comply with Sections (B)(2), (B)(3)(a), and (B)(3)(d) under *Section B. Nationally Covered Indications* for NCD 20.32
I	90702621 Rev/Ver AF	Section 4.5	"Small sites" will be - removed for poolability analyses	Update to reflect the analysis to be performed (match the Section 5.5.1 from FDA recommendation).
I	90702621 Rev/Ver AF	Section 5.6	Model covariates	Pre-specified covariates for the univariate and multivariate models
I	90702621 Rev/Ver AF	Section 7.3.2	Adverse event dates are "Onset date", "If Hospitalization, is this a new hospitalization, If Yes, admission date", and "If Yes, discharge date" from the Adverse Event Form	Add new hospitalization admission date and discharge date as valid dates for Days to Last Follow-up.