STUDY CODE: TPS003 **STUDY NAME: PERSIST-AVR**

STUDY TITLE:

Perceval Sutureless Implant Vs Standard Aortic Valve Replacement

STATISTICAL ANALYSIS PLAN

Version N° 5.00 **Date: October 8th , 2019**

Confidentiality Statement

Part or all of the information in this statistical analysis plan may be unpublished material. Accordingly, this document is to be treated as confidential and restricted to its intended use. This material is the property of LivaNova PLC and must not be disclosed or used except as authorized in writing by LivaNova PLC.

	Name	*Signature*	*Date*
Prepared by	Study Biostatistician	Moi Jiang	09OCT2019
Approved by	Clinical Project Manager		09-OCT-2019
Approved by	Director, Statistics and Data Management	Giacomo Mordenti	09 Oct 2019

STUDY CODE: TPS003 **STUDY NAME: PERSIST-AVR**

STUDY TITLE:

Perceval Sutureless Implant Vs Standard Aortic Valve Replacement

STATISTICAL ANALYSIS PLAN

Version N° 5.00 **Date: October 8th , 2019**

	Name	*Signature*	*Date*
Prepared by	Study Biostatistician		
Approved by	Clinical Project Manager		
Approved by	Director, Statistics and Data Management		

TABLE OF CONTENT

1. INTRODUCTION

This statistical analysis plan (SAP) includes all definitions and analysis details for the analysis of the study TPS003 PERSIST-AVR in accordance with the amended Clinical Investigational Plan (CIP) dated 27FEB2018, and the e-CRF version MSC007 dated 02MAY2018. The analysis will be performed by the Department of Global Biometrics at LivaNova in accordance with this SAP with the support of a Independent Statistical Unit (ISU) for the planned interim analyses for sample size selection.

2. STUDY DESIGN

2.1. STUDY OBJECTIVES

2.1.1. Primary objective

The primary objective is to test the safety and efficacy of Perceval valve versus standard sutured stented bioprosthetic aortic valves among the intended trial population (subjects with severe symptomatic aortic stenosis or steno-insufficiency who are candidates for surgical replacement of their native aortic valve).

2.1.2. Secondary objectives

The secondary objectives are to assess effectiveness, safety, and hemodynamic performance of Perceval valve as compared to standard sutured stented bioprosthetic aortic valves.

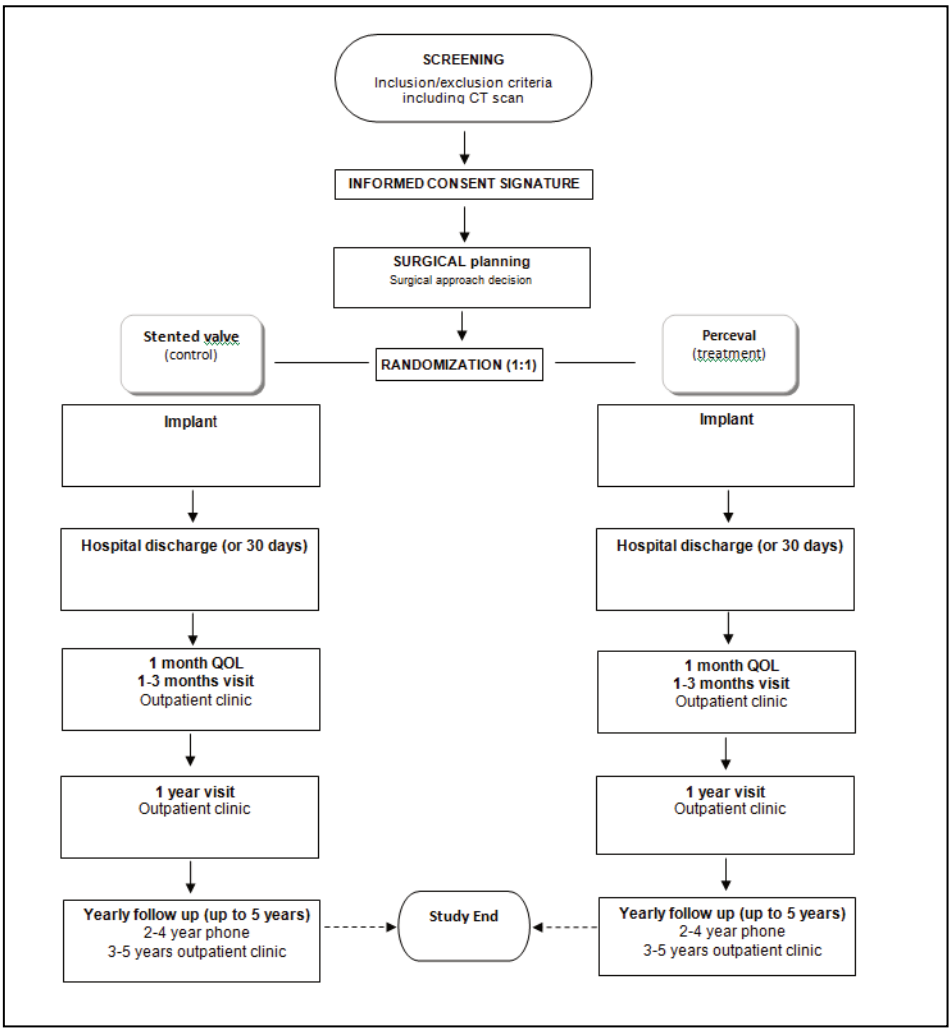
2.2. OVERALL STUDY PLAN

PERSIST-AVR is a prospective, randomized (1:1), adaptive, open, multi-center, international study.

The trial plans for an inclusion period of approximately 30 months at approximately 60 international sites where the Perceval is commercially available. Subjects will be followed for 5 years. Total study duration of approximately 7.5 years is expected. Subjects will be evaluated at screening (the preoperative assessments will be considered as baseline), at discharge visit following prosthesis implant, between one and three months, at 1 year, and annually up to 5 years.

Study flowchart is graphically reported in Figure 1.

Figure 1: Study Flowchart



2.3. SAMPLE SIZE CALCULATION

Total sample size for this trial will be adaptively determined through interim analyses. The sample size will be approximately 900, 1050, or 1234 subjects.

The sample size is based on the analysis of primary endpoint (section 10.1) for assessing non-inferiority of Perceval valve arm compared to the control.

Based on CAVALIER experience, propensity score matched studies[1,4,5,6] and, GARY registry[7] the freedom from MACCE composite endpoint is expected to be similar for both arms. The original sample size of 1234 patients was based on CAVALIER experience which reported a freedom from MACCE at 1 year equal to 88%. Blinded review of the data for the ongoing current trial (database snapshot December 2017) estimated the pooled one year MACCE-free rate across arms at 91%, showing an overall primary endpoint higher than originally assumed. Due to this reason, 1234 patients may lead to an overpowered study and a smaller sample size may be required to demonstrate study objectives. In

order to adapt the sample size to the real study performance an adaptive design has been included in the study protocol.

Table 1 summarizes the overall operating characteristics of the design over a range of assumptions for the true underlying MACCE-free rate on control ($MACCE_{ctrl}$) and treatment ($MACCE_{PERCEVAL}$). The information provided by the table are:

- Pr(success): it is the proportion of simulated trials achieving success at the final analysis.
  - when the true effect is exactly at the non-inferiority margin (-0.05), it represents the Type I error; otherwise
  - Pr(success) is the power.
- E[N]: it represents the mean sample size across simulated trials, reflecting an average over some trials that stop with N = 900, some at N = 1050, and others at N = 1234.

Table 1 : Overall operating characteristics

True $MACCE_{ctrl}$	*True* $MACCE_{PERCEVAL}$	*True Effect*	*Pr(success)*	*E[N]*
0.88	0.83	−0.05	0.025	1228.7
	0.88	0.00	0.704	1148.9
0.91	0.86	−0.05	0.023	1228.8
	0.91	0.00	0.805	1128.7

Simulations indicate that the adaptive design is able to control the Type I error and Power of the study while looking for the most appropriate sample size. In fact, under the expected assumption of equivalent arms with 91% MACCE-free rate, the trial has approximately 80% power, the type I error rate is controlled to 2.5% one-side and the average sample size is reduced to approximately 1129.

The adaptive design is fully described in CIP Appendix 8 along with additional summaries of the operating characteristics.

2.4. RANDOMIZATION

Subjects are randomly assigned in equal numbers (1:1) to one of two treatment arms:

- Perceval valve
- Standard sutured stented bioprosthetic aortic valve (control)

on the basis of a randomization list that has been performed before study start. Randomization list is stratified by 2 factors:

- country, and
- surgical approach (full sternotomy" or "mini-sternotomy").

Only subjects who have undergone standard chest CT-scan to determine if the aortic stenosis can be replaced with an available Perceval valve, and is potentially suitable for mini-sternotomy, may be randomized. For each subject the randomization will be performed through eDC system after enrollment and before valve operation, specifically after CT scan and surgical approach ("full sternotomy" or "mini-sternotomy") has been determined.

One patient will be considered adherent to the randomization if an attempt to implant of the allocated valve is performed. In case of, after an unsuccessful implant, a valve of the other arm is implanted the adherence to randomization allocation will be considered maintained.

3. DOCUMENT AND CHANGE HISTORY

3.1. CHANGES IN ANALYSIS COMPARED TO CLINICAL INVESTIGATIONAL PLAN

No changes in the analysis as described in the Clinical Investigational Plan.

3.2. SAP AMENDMENT RATIONALE AND CHANGE HISTORY

Table 2 : SAP Amendment description

Version	*Date*	*Section(s)*	*Description of modifications*
1.0	13MAR2017		Initial Release.
2.0	26JUL2017	7,9,10.10,13	Additional general details on how to represent values in the statistical analysis. Additional specifications have been added to analysis populations (including the definition of adherence to randomization allocation). Details on compliance calculation have been added. Additional derivation rules.
3.0	28JUN2018	10, 11, 12 Editorial changes impact all the document	SAP has been amended following the CIP amendment: • Adaptive design; • Primary analysis; • Sensitivity analysis. Details on primary analysis (Bayesian approach) have been added. Additional analysis for NYHA, echocardiographic data, EQ5D. Surgical times analysis has been modified from ANOVA to ANCOVA. Editorial changes (new template per internal SOP). List of TFLs have been removed (included in a different document).
4.0	16MAY2019	12.1.2, 12.2, 12.3, 13.4	Sensitivity analyses of Surgical times have been added; Additional details for checking the assumptions for ANCOVA; Additional details have been added for the calculation and testing of NCI and rNCI, including the management of missing data (if needed) and an analysis by country; Sensitivity analyses of EQ-5D-5L have been added (using France-based scoring system and with a Mann-Whitney test); The use of a discount rate for the QALY has been removed; Additional details on calculation of EQ-VAS have been added; Details on pacemaker evaluation through ECG Core Lab have been added.

5.0	01OCT2019	6.6, 8, 8.6, 11, 13.1.1, 13.1.2, 18.10, 18.12	Usage of analysis populations have been updated. Statistical tests and p-value have been added for comparison of baseline variables between treatment groups. Site reported echocardiographic parameters have been updated. Core-lab reported echocardiographic parameters have been updated. In Cox regression model of MACCE, language has been modified to reflect that for the stratification factor of country, countries with less than 20 enrolled patients will be polled. Summary of the number and percentage of subjects with AEs has been updated. Summary of the number and percentage of TEAEs has been updated. "Parivalvular leak" has been updated to be "central leak". Additional languages have been added for instructions on how to handle missing AE start date and stop date. Data derivations have been updated.

4. OVERVIEW OF PLANNED STATISTICAL ANALYSIS

The study plans for the different statistical analyses. The statistical analyses will be performed as detailed in this Statistical Analyses Plan (SAP). The SAP will be finalized prior to any interim analyses.

4.1. PRIMARY STATISTICAL ANALYSIS

The primary statistical analysis will be performed by Sponsor personnel once all the subjects complete their 12 month follow-up and the study database will be locked. This analysis will be focused on:

- primary endpoint
- mortality and morbidity rates up to 365 days
- other safety endpoints up to 12-months FU visit

The results of the primary statistical analysis will be the basis for a dedicated clinical study report.

4.2. INTERIM STATISTICAL ANALYSES

Interim analyses for accrual stopping will be conducted by an Independent Statistical Unit (ISU) composed by external statistical personnel not involved in the study conduct at the time when approximately 900 and 1050 patients are enrolled in the study. For details on the interim analysis please refer to Section 15.

4.3. OTHER STATISTICAL ANALYSIS

Safety monitoring analyses based on unclean data will be performed periodically with the involvement of Data Safety Monitoring Board (DSMB). The analysis will be descriptive in nature and no tests on the primary endpoint will be performed. Therefore no alpha adjustment for multiplicity is required.

The DSMB will be blinded to the actual treatment arms and will be responsible for monitoring the safety and well-being of the subjects participating in the trial, ensuring the scientific integrity of the trial, and recommending action items based on safety issues including trial termination as warranted. DSMB can be unblinded at any time during the study for safety reasons.

The first Safety monitoring analysis will be performed after first 200 subjects are enrolled. Subsequent analyses will be carried out approximately every 200 enrolled subjects.

4.4. FOLLOW-UP AND FINAL STATISTICAL ANALYSES

Follow-up statistical analyses will be performed by Sponsor personnel once per year once all the subjects complete their 2, 3, 4, and 5 years follow-up and the study database will be locked. The 5-year analysis will be considered as the final study analysis and will be the basis for the integrated clinical study report.

4.5. ADDITIONAL STATISTICAL ANALYSIS

Potential statistical analyses might be performed after Primary statistical analysis upon scientific purpose. Analyses will be detailed through an amendment to the SAP or a specific document.

Specific SAE listings that will contain all information in the SAE CRF needed to screen for CEC adjudicability will be created. These listings will be provided to the Safety Office and CPM upon requests and include:

- List of cases with CKMB/Troponin elevations as described in the adjudicability criteria
- List of EKG tracing results per visit
- List of mRS and NIHSS results for each patient per visit
- List of hemoglobin/hct; platelet count results; CRP results

5. ANALYSIS CONVENTION

5.1. GENERAL PRINCIPLES

The statistical analysis will be performed on the analysis study database with SAS version 9.4 or above (SAS Institute, Cary, N.C.), FACTS version 6.1 or above (Berry Consultants, Austin TX) or R version 3.3 or above.

All tabular and graphical presentations will be done by treatment group.

All the data collected and derived in the study will be presented in subject data listings.

5.2. DESCRIPTIVE STATISTICS

Descriptive statistics will be calculated using as reference the number of subjects in the relevant analysis population (any exception will be specified) according to the nature of the data as follows:

- Continuous variables: number of non-missing observations, arithmetic mean, standard deviation, minimum and maximum values, median and quartiles. If there are less than 5 observations, only the number of non-missing observations, arithmetic mean, median, minimum and maximum will be presented.
- Categorical variables: Number of missing and non-missing observations the relevant percentage on the analysis population. In case of subcategories, the relative frequencies will be calculated on the basis of the subjects in the subcategory, in this case a footnote will be added explaining the different denominators.
- Time-to-event variables: number of non-missing observations, minimum, first quartile, median including 95% confidence interval, third quartile, and maximum. For calculating the survival estimate CI bounds the log-log transformed estimate of CI bounds will be used. In addition, Kaplan-Meier estimates and plots will be provided with:
    - the 95% confidence interval bounds calculated per the method proposed by Greenwood8
    - the respective number of patients at risk and Kaplan-Meier estimates at different time points (at least every 5 months or more frequently)
    - the median and its 95% confidence interval.
    - the log-rank test p-value
    - the Hazard Ratio (HR), its 95% confidence interval and the relevant p-value.

  Subjects ongoing and who are free from event at the analysis cut-off date will be censored at the analysis cut-off date. Subjects who have discontinued without an event will be censored at the date of discontinuation.
- For Bayesian analyses: posterior means, standard deviations, and 95% credible intervals will be reported.

5.3. SUBGROUPS DEFINITIONS

The following variables and levels will be considered for subgroup analyses:

- Gender (“Female” vs “Male”)
- Surgical approach (“Full sternotomy” vs “Mini-sternotomy”)
- Concomitant procedures (“Concomitant procedure” vs “Isolated procedure”)
- STS predicted risk of mortality (“Low risk [<4]” vs “Med/High risk [≥4]”)
- Regurgitation severity (“None/Trace” vs “Mild/Moderate/Severe”) at discharge
- Early (≤ 30 days) Post-implant permanent PaceMaker (“Yes” vs “No”)

5.4. DEFINITION

Table 3 : Definitions

Term	Definition
Investigational Device (ID)	Perceval valve is a bioprosthetic heart valve made of bovine pericardium, stabilized in a buffered glutaraldehyde solution and assembled on a nitinol stent. The device is indicated for the replacement of a damaged or malfunctioning native heart valve in humans via traditional surgery.
Study Day Count	The day of implant is defined as study Day 1. Calculate the study day according to the following rules: If date < study Day 1 then study day = Date – study Day 1 If date > study Day 1 then study day = Date – study Day 1 +1
Baseline	Last preoperative assessment
Screening Failure	Screened patients who do not fulfill IE criteria OR patient who signed the IC but withdrawn from the study for termination reason is “Failure to meet randomization criteria”
Study completers	Patients who complete 5 year visit. The termination reason is “Completed the Study”
As-treated principle	Patients will be assigned to the treatment groups according to the device actually implanted to the subject.
As-randomized principle	Patients will be assigned to the treatment groups according to the randomization list independently from the actual implant to the device under investigation.

Term	Definition
Study phases	• Screening & enrolment • Operative evaluation • Hospital Discharge Assessments at: • 1 month (M1): 30 days ± 7 days. The information (EQ5D) should be submitted to the site during the 1-3 months visit. • 1 to 3 months FU visit (M3): between 30 and 90 days • 1 year FU visit (Y1): 1 year (365 days) ± 30 days after the day of the surgery • 2 years FU visit (Y2): 2 years (730 days) ± 30 days after the day of the surgery • 3 years FU visit (Y3): 3 years (1095 days) ± 30 days after the day of the surgery • 4 years FU visit (Y4): 4 years (1460 days) ± 30 days months after the day of the surgery • 5 years FU visit (Y5): 5 years (1825 days) ± 30 days after the day of the surgery

6. ANALYSIS POPULATION

6.1. ENROLLED ANALYSIS POPULATION

The "Enrolled population" (ENR) is defined as all patients who will be included in the study (enroll date but no screening failure) with a signed and dated informed consent. The ENR population will be used mainly for subject disposition summaries.

6.2. SAFETY ANALYSIS POPULATION

The "Safety" population (SAF) is defined as all randomized and implanted patients. The SAF population will be used mainly for safety summaries. Patients will be allocated to the treatment groups according to the «as-treated» principle.

6.3. INTENTION-TO-TREAT ANALYSIS POPULATION

The "Intention-To-Treat" population (ITT) is defined as all randomized patients. Patients will be allocated to the treatment groups according to the «as-randomized» principle.

6.4. MODIFIED INTENTION-TO-TREAT ANALYSIS POPULATION

The "modified Intention-To-Treat" population (mITT) is defined as all randomized patients (ITT population) who receive a study valve (Perceval or standard sutured stented valve). Patients will be allocated to the treatment groups according to the «as-randomized» principle. The mITT population will be used as sensitivity efficacy analyses.

6.5. PER-PROTOCOL POPULATON

The "Per-Protocol" population (PP) is defined as all randomized and implanted patients attending the scheduled follow-up visits up to 1 year at minimum and without the following major protocol deviations:

- Assessment/Procedure (related to primary endpoint) Incomplete/Not Done
- Serious Adverse Event Not Reported
- No adherence to randomization allocation
- No consent obtained or if consent obtained after implant.

Further major protocol deviation may be detected by the Steering Committee and appropriately reported in the eCRF. In the exceptional event that "other" reasons requiring exclusion from the PP become apparent during the final data review, these will be suitably documented at that time in the final data review report (which will be prior to database lock).

Primary efficacy analysis will be performed on this population. The PP population will be used for analyses of primary and secondary endpoints.

6.6. USAGE OF ANALYSIS POPULATIONS

Table 4 : Definitions Use of analysis sets

	ENR	*SAF*	*mITT*	*PP*
Subject disposition	X			
Subject discontinuation	X		X	
Protocol deviations			X	
Subject demographics		X	X	X
Implant characteristics		X	X	X
Valve characteristics		X	X	X
CT Scan		X	X	X
Medical history		X	X	X
Other baseline characteristics		X	X	X
Healthcare Economics		X	X	
Compliance			X	X
Primary endpoint			X	X
Secondary efficacy endpoint		X	X	X
Secondary Cost effectiveness Analysis		X	X	
Analysis of Adverse Events	X	X	X	
Other safety parameter		X	X	

7. DISPOSITION

7.1. SUBJECT DISPOSITION

All presentations for subject disposition will be by treatment group, and overall.

For describing the subject disposition, the following populations will be summarized:

- Subjects enrolled and eligible (overall)
- Subjects eligible but not randomized and reason for non-allocation (overall)
- Subjects randomized
- Subjects implanted, implanted with the planned valve, who crossed the arm, implanted with a non-study device, not implanted
- Subjects in each analysis population
- Subjects on-trial
- Trial completers
- Subjects who discontinued the trial prematurely (and reason).

For the overall report, the percentage denominator will be the number of enrolled subjects. For the "by treatment group" calculations, the percentage denominator will be the number of randomized subjects within each arm.

7.2. SUBJECT DISCONTINUATION

If more than 30% of randomized subjects discontinue the study the distribution of the time to discontinuation from study will be summarized using time-to-event methods (section 5.2). The time from randomization to discontinuation will be censored at the End-of-Study Visit or Analysis cut-off date.

In addition, the following information will be provided descriptively summarized:

- Study length
- Extent of exposure and total patient-years

7.3. PROTOCOL DEVIATIONS

All protocol deviations (PDs) will be summarized overall (ENR) and by arm (ITT) by severity (minor and major) and grouped into different categories:

- Inclusion/Exclusion
- Informed Consent
- Procedure Performed by Unauthorized Physician
- Visit/Follow-up Not Done

- Visit/Follow-up Out of Window
- Assessment/Procedure Incomplete/Not Done
- Assessment/Procedure Out of Window
- Planned Surgical Approach Changed
- Randomization
- Study Device Not Used per IFU or per Protocol
- Serious Adverse Event Not Reported per Protocol
- Other

PDs will be reported as:

- total PDs on the total number of patients
- number of subjects with at least one PD on the total number of patients

Protocol deviations will be presented in a subject data listing.

8. DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

The demographics/baseline characteristics will be analyzed by means of summary statistics (section 5.2) according to the nature of the variable (continue or discrete). Statistical tests for comparison of demographic and baseline characteristics between treatment groups will be performed according to the nature of the variable. A total of around 100 demographic and baseline variables will be compared, therefore a Bonferroni correction will be applied in order to take into account the multiple comparison. Given the correction, the distribution of the variable will be considered statistically significant different between treatment groups if the p-value is <0.0005.

8.1. SUBJECT DEMOGRAPHICS

Subject demographics are country, age (years), gender, weight [kg], height [m], body mass index (BMI) [kg/m2], Society of Thoracic Surgeon (STS) Predicted Risk of Mortality [%], and Logistic EuroScore II [%].

8.2. IMPLANT CHARACTERISTICS

The implant characteristics will include:

- Surgical details (including, but not limited to, Surgical Approach, Cannulation, Debridement, Universal sizer measurement, bicuspidy, and decalcification)
- Implant Concomitant Procedures
- Surgical Time (section 18.12)
    - Operating room suite (min)
    - Skin-to-skin (min)

  - Extracorporeal circulation (min)
  - Cross clamp (min)
- Post Implant Intraoperative Transesophageal Echocardiogram
- Cardiac Rhythm-Post Operative Evaluation
- Intraoperative Management

8.3. VALVE CHARACTERISTICS

The following variables collected in the 'Valve implant' CRF will be analyzed by means of summary statistics:

- Valve manufacturer
- Number of valves used (in cases of failure to implant)
- Valve model
  - If Perceval: Pass of Perceval sizer obturator through the annulus, Labeled valve size implanted, Removal of the guiding suture, post-dilatation using ballooning
  - If conventional valve: Valve sizing measured with manufacturer sizers, Labeled valve size implanted, Valve positioning, Suture technique, use of pledget and Cor-Knot
- Implantation Outcome
  - Implantation Outcome
  - Implantation of the randomized valve

8.4. CT SCAN

The characteristics of native valves measured preoperatively by CT scan are: LVOT diameter, Aortic Annulus mean diameter, Annulus perimeter, Annulus surface, Sinus of Valsalva diameter, Sino tubular junction (STJ) diameter, Ascending aorta, and Ratio STJ/annulus.

8.5. MEDICAL HISTORY

Medical history and patient risk factors will include Coronary Artery Disease, Angina, Myocardial Infarction, Heart Failure, Carotid Artery disease, Previous Carotid Artery Surgery or Percutaneous Transluminal Carotid Angioplasty (PTCA), Stroke, Transient Ischemic Attack (TIA), Endocarditis, Peripheral Vascular Disease, Pulmonary Hypertension, Chronic Lung Disease, Diabetes, and Dyslipidemiapecify.

8.6. OTHER BASELINE CHARACTERISTICS

For parameters collected on more than 1 occasion during the study including baseline, the assessment at baseline will be presented with assessments collected later on in the study and not in a separate table. These parameters are:

- NYHA class
- Signs of Infections
- Modified Rankin Scale
- NIHSS Score
- Vital signs
- Electrocardiogram (ECG)
- Site Reported Echocardiographic Values:
    - Aortic Mean Pressure Gradient (MPG), Aortic Peak Pressure Gradient (PPG), Sub aortic gradient, Left Ventricular Ejection Fraction (LVEF), Effective Orifice Area (EOA), Effective orifice area indexed (EOAi), Aortic Peak Velocity (VmaxAORTIC), LVOT Velocity max (VmaxLVOT), Doppler velocity index, LV mass, and Septum thickness.
    - Paravalvular and Central Leaks
    - MPG classification (section 18.12)
- Core-lab Reported Echocardiographic Values:
    - Aortic Mean Pressure Gradient (MPG), Aortic Peak Pressure Gradient (PPG), Mean Pressure Gradient, LVOT Peak Pressure Gradient, Mean Gradient (P2-P1), Peak Gradient 4(V2A- V2L) Left Ventricular Ejection Fraction (LVEF), Effective Orifice Area (EOA), Effective orifice area index (EOAi), Left Ventricular Mass, Left Ventricular Mass Index, Left Ventricular End Diastolic Volume, Left Ventricular End Systolic Volume, Aortic Peak Velocity (VmaxAORTIC), LVOT Velocity max (VmaxLVOT), Cardiac Output, and Cardiac Index.
    - Paravalvular and Central Leaks
- Health status (EQ-5D-5L)

  For each time point, a radar chart will be created representing on the radii the mean level of each dimension.
- Blood test
    - Blood test results
    - Clinical significance of blood test values

NYHA, Echocardiographic values, Health Status, and Blood test will be analyzed to assess the trend during time. Additional details in section 11.1, 11.2, 0, and 13.2 respectively.

8.7. PRIOR AND CONCOMITANT CARDIOVASCULAR MEDICATIONS

Prior and concomitant cardiovascular medications are collected in the e-CRF as per enrollment. For the analysis, the following algorithm will be used to define prior and concomitant medication:

- Prior is all medication stopped prior to study day 1, regardless of its start date.

- Concomitant is any medication not stopped before the implant of the valve, regardless of its start date or medication started after the implant of ID.

Medication will be summarized and sorted alphabetically separately for prior and concomitant medication by Medication Type.

For each medication, the number of subjects will be displayed.

8.8. HEALTHCARE ECONOMICS

The healthcare economics data are:

- Intraoperative Management parameters (Total Number of Valve(s) Used, Cardioplegia, Bloods Units, Fluids, Inotropes, Coagulation Factor, and Hemostatic Agent), and
- Intensive Care Unit (ICU) Management data (Ventilation Time, Loss of Bleeding, Level of Lactate, Level of Creatinine during ICU, Renal replacement therapy, Use of inotropes, transfusion during ICU and details of plasma, platelet and red blood cell numbers).

9. COMPLIANCE

Compliance will be summarized to assess adherence to protocol requirements on:

- FU visits,
- NYHA class assessment,
- Site Echocardiograpy,
- Core Lab Echocardiograpy,
- ECG and
- Blood tests.

Compliance will be defined as the number of visits/assessments actually done by a subject divided by the number of visits/assessments planned for the respective period and multiplied by 100:

V_i Compliance [%] = [# assessments at V_i / planned # assessments at V_i] x 100.

An assessment will be considered as performed if a complete “assessment date” is available in the corresponding CRF.

The number of planned assessments by visits are:

- 1 to 3 months FU visit:
  - Number of patients who perform 1-3 months visit/assessment +
  - Number of patients who don’t perform 1-3 months visit/assessment with >90 days FU post-implant
- 1 year FU visit:
  - Number of patients who perform 1 year visit/assessment +

  - Number of patients who don't perform 1 year visit/assessment with >395 (365+30) days FU post-implant
- 2 years FU visit:
  - Number of patients who perform 2 years visit/assessment +
  - Number of patients who don't perform 2 years visit/assessment with >760 (730+30) days FU post-implant
- 3 years FU visit:
  - Number of patients who perform 3 years visit/assessment +
  - Number of patients who don't perform 3 years visit/assessment with >1125 (1096+30) days FU post-implant
- 4 years FU visit:
  - Number of patients who perform 4 years visit/assessment +
  - Number of patients who don't perform 4 years visit/assessment with >1491 (1461+30) days FU post-implant
- 5 years FU visit:
  - Number of patients who perform 5 years visit/assessment +
  - Number of patients who don't perform 5 years visit/assessment with >1856 (1826+30) days FU post-implant

At each time point, patients who complete the visit/assessment will be classified as "Completed inside window" according to the following rules:

- 1 to 3 months FU visit: if visit/assessment is perfomed within 30-90 days after the day of the surgery
- 1 year FU visit: if visit/assessment is performed at 365 ± 30 days after the day of the surgery
- 2 years FU visit: if visit/assessment is performed at 730 ± 30 days after the day of the surgery
- 3 years FU visit: if visit/assessment is performed at 1095 ± 30 days after the day of the surgery
- 4 years FU visit: if visit/assessment is performed at 1460 ± 30 days after the day of the surgery
- 5 years FU visit: if visit/assessment is performed at 1825 ± 30 days after the day of the surgery

Patients who complete the visit/assessment not classified as "Completed inside window" will be classified as "Completed outside window".

10. PRIMARY ANALYSIS

10.1. PRIMARY ENDPOINT

The primary endpoint to assess the safety and efficacy of Perceval valve versus standard sutured stented valves is the freedom from MACCE (composite endpoint of all cause death, myocardial infarction, stroke, and valve re-intervention) at one year based on CEC adjudication.

10.2. TESTING STRATEGY AND MULTIPLICITY ADJUSTMENT

The primary analysis will compare Perceval valve (Treatment arm) vs. standard sutured stented valve (Control arm) on the Per Protocol population.

A one-sided non-inferiority test, using the non-inferiority margin Δ=0.05, will be performed to compare the two arms on the proportion of subjects that are event-free at one year (primary analysis). The null hypothesis is that the Perceval arm is inferior to the Control arm in the primary endpoint:

H_0: $MACCE_{CONTROL} - MACCE_{PERCEVAL} \geq \Delta$

H_1: $MACCE_{CONTROL} - MACCE_{PERCEVAL} < \Delta$

where $MACCE_j$ is the one-year event-free rate for arm j.

Superiority will be tested if non-inferiority is met. In this case, the null hypothesis is that the Perceval arm is not superior to the Control arm in the primary endpoint:

H_0: $MACCE_{PERCEVAL} \leq MACCE_{CONTROL}$

H_1: $MACCE_{PERCEVAL} > MACCE_{CONTROL}$

where $MACCE_j$ is the one-year event-free rate for arm j.

This superiority test is nested hierarchically within the non-inferiority comparison and is a closed-testing procedure, thus no alpha adjustment is required.

10.3. PRIMARY ANALYSIS

The primary analysis will calculate the Bayesian posterior probability that the difference [$MACCE_{CONTROL} - MACCE_{PERCEVAL}$] is lower than 0.05 (predetermined non-inferiority margin):

$$Pr(MACCE_{CONTROL} - MACCE_{PERCEVAL} < 0.05 \mid Data)$$

The null hypothesis will be rejected, and non-inferiority concluded, if the posterior probability exceeds 0.9775 at the final analysis. This threshold was selected based on simulations to control the Type I error rate at 2.5% one-sided (CIP Appendix 8).

The probability $Pr(MACCE_{CONTROL} - MACCE_{PERCEVAL} < 0.05 \mid Data)$ is calculated on the basis of simulation after the definition of the posterior distribution of p_j according to the following steps:

1. The number of responders X_j is modeled in arm j as $X_j \sim \text{Binomial}(n_j, p_j)$ where n_j is the number of subjects for arm j and p_j is the proportion of subjects that are event-free at one year (response rate);
2. The response rates p_j are then transformed and modeled independently on the log-odds scale. They have a Normal prior distribution:

$$[\theta_j] = \left[\log\left(\frac{p_j}{1-p_j}\right)\right] \sim N(\theta_0, \sigma_0^2)$$

with $\theta_0 = 1.99$ and $\sigma_0^2 = 3^2$ so that when converted back to the probability scale, this prior is centered at 0.88 (the expected event-free rate for the control arm);

3. One million (1000000) posterior samples for p_j in each arm will be drawn using a standard Markov chain Monte Carlo (MCMC) technique (random walk Metropolis-Hastings alogorithm) after discarding the initial set of burnin samples. The calculation of the probability Pr(MACCE$_{CONTROL}$ – MACCE$_{PERCEVAL}$ < 0.05 | Data) will be done as

$$\frac{\#\text{times } [\text{MACCE}_{\text{CONTROL}} - \text{MACCE}_{\text{PERCEVAL}} < 0.05]}{1000000}$$

In case of test for superiority, the Bayesian posterior probability of superiority

Pr(MACCE$_{CONTROL}$ < MACCE$_{PERCEVAL}$ | Data)

will be assessed in the Per Protocol population and will be provided. Superiority will be claimed if the above probability exceeds 0.9775. In order to calculate the probability of superiority, one million (1000000) simulations will be run using the posterior distribution on p_j in each arm, and the calculation of the probability Pr(MACCE$_{CONTROL}$ < MACCE$_{PERCEVAL}$ | Data) will be done as

$$\frac{\#\text{times } [\text{MACCE}_{\text{CONTROL}} < \text{MACCE}_{\text{PERCEVAL}}]}{1000000}$$

10.4. SENSITIVITY ANALYSES

The primary analysis will be assessed in sensitivity analyses using different assumptions and populations:

- the primary analysis will be repeated on the mITT population. For this analysis, missing data for the primary endpoint will be imputed using the same procedure used for the interim analysis (see Section 15). This analysis will be performed to assess the robustness of the results to protocol deviations: mITT analysis provides a “real life” scenario, in which participants may not comply to the protocol.
- the proportion of subjects that are event-free at one year will be estimated by means of a logistic regression model including as covariates the implanted valve groups (Perceval or standard sutured stented valves) and the stratification factors:

- country, and
- surgical approach.

This method allows to assess the sensitivity of non-inferiority results to possible confounding impact of the stratification factors.

A 97.5% one-sided upper confidence limit will be computed from the logistic model for the difference ($MACCE_{CONTROL}$ – $MACCE_{PERCEVAL}$). The non-inferiority of Perceval to the Control will be confirmed if the upper confidence limit is lower than 0.05 (Δ). This sensitivity analysis of non-inferiority will be assessed in the Per Protocol Population.

- the one-year freedom from MACCE will be evaluated using the Kaplan-Meier estimates for cumulative freedom from MACCE and the Greenwood standard errors for each arm (as planned in the CIP before amendment). A 97.5% one-sided upper confidence limit will be computed for the difference ($MACCE_{CONTROL}$ – $MACCE_{PERCEVAL}$) where $MACCE_j$ is the survival estimated by the Kaplan-Meier algorithm for arm j. The Perceval arm will be confirmed not inferior to the Control if the upper confidence limit is lower than 0.05 (Δ).

  The test statistic is

$$\frac{\mathrm{MACCE_{PERCEVAL}} - \mathrm{MACCE_{CONTROL}} + \Delta}{\sqrt{\mathrm{VAR(MACCE_{PERCEVAL})} + \mathrm{VAR(MACCE_{CONTROL})}}}$$

  where VAR()s are the variances estimated by Greenwood's formula. The null hypothesis will be rejected if the test statistic is greater than 1.96. This sensitivity analysis of non-inferiority will be assessed both in PP and mITT populations in order to confirm the robustness of results to statistical method and protocol deviations.

11. SECONDARY EFFICACY ANALYSES

The following secondary efficacy analyses will be performed:

- each component of the primary endpoint will be analyzed as binomial proportion (95% C.I.) of subjects event-free at one year (section 18.12);
- primary endpoint will be analyzed annually as binomial proportion (95% C.I.) of subjects event-free at 1, 2, 3, 4 and 5 years of follow-up (section 18.12);
- The time to the first MACCE will be analyzed using all data collected up to analysis cutoff date (patients may have more that 1-year follow-up):
  - by means of time-to-event methods (see section 5.2)
  - with a Cox regression model, using the stratification factors (country and surgical approach) and pre-operative variables (including demographic, medical history, and concomitant procedures) as covariates. Countries with less than 20 enrolled patients will be pooled. The assumptions of proportionality will also be investigated with a time-dependent exploratory variable, which is defined as treatment*{log[time to event]}. If the p-value from the Wald Chi-squared statistic for this variable is less than 5% there is

evidence of a departure from the adjusted model assumptions. In case of non-proportional hazards, a Weibull model will be assumed as parametric form of the distribution of survival times. The treatment group will be included in the model as covariate.

11.1. NYHA

The trend of NYHA classes during time will be assessed by means of:

- shift tables from preoperative to annual post-operative FU assessments (from 1 year FU to 5 year FU) by treatment group;
- a generalized linear model for longitudinal data on ordinal scale in order to model the probabilities of having lower NYHA response levels during time. The model will include the treatment covariate and the visit assessment as covariates, and the response categories for the NYHA will be considered in following order: Class I, Class II, Class III, and Class IV. An unstructured correlation structure will be defined in order to allow the model to take into account the dependency of repeated observations within each subject. The odds ratio of Perceval being in lower NYHA categories compared Standard valve will be shown.

11.2. SITE REPORTED ECHOCARDIOGRAPHIC DATA

11.2.1. Hemodynamic data

Mean changes from baseline in site-reported valve hemodynamic

- mean gradient,
- peak gradient and
- LV mass

will be analyzed based on data observed while the subject remains on study (observed cases) as well as data imputed using multiple imputation (MI) methodology for time points at which no value is observed.

Multiple imputation will be performed under the assumption of missing-at-random (MAR) and will be implemented in two steps:

- First, partial imputation assuming MAR will be carried out to impute intermittent (non-monotone) missing data based on a multivariate joint Gaussian imputation model using the Markov chain Monte Carlo (MCMC) method. A separate imputation model will be used for each treatment arm. The imputation models will include countries, surgical approach, gender, age, hemodynamic assessments at each time point from Baseline. The MCMC method will be used with multiple chains, 200 burn-in iterations, and a non-informative prior. In case of non-convergence or non-estimability issues, a single model will be considered with treatment arm added as explanatory variable to the model.
- The remaining monotone missing data will be imputed using sequential regression multiple imputation, where a separate regression model is estimated for imputation of each variable

(i.e., measurement at each time point). Each regression model will include explanatory variables for country, surgical approach, gender, age, treatment and all previous values of hemodynamic assessment from baseline.

No rounding or range restrictions will be applied to imputed continuous values.

Imputed data will consist of 200 imputed datasets. The random seed number for partial imputation with the MCMC method will be 200582, and the random seed number for the sequential regression multiple imputation will be 4112.

Each of the 200 imputed datasets will be analyzed using observed and imputed data.

Change in hemodynamic assessment from baseline to 1-year (5-year) visit will analyzed by means of a mixed regression model analysis including the treatment, hemodynamic assessment at baseline, country, surgical approach, gender, and age.

Treatment group comparison at Visit T will be based on the least squares mean (LSM) difference between treatment groups in change from baseline in hemodynamic assessment estimated by the analysis model in each of the imputed datasets. Results from analysis of each imputed dataset, i.e., LSM treatment differences and their standard errors, will be combined using Rubin's imputation rules to produce a pooled LSM estimate of treatment difference, its 95% confidence interval, and a pooled p-value for the test of null hypothesis of no treatment effect.

11.2.2. **Paravalvular and Central Regurgitations**

Paravalvular and Central leaks severity will be classified in

- none/trace,
- mild and
- moderate/severe

(section 18.12) and will be analyzed based on data observed while the subject remains on study as well as data imputed using multiple imputation (MI) methodology for time points at which no value is observed.

Multiple imputation will be performed under the assumption of missing-at-random (MAR) and will be implemented in two steps:

- Leak severity variable (categorical) will be coded with binary dummy indicator variables for the inclusion in the imputation model.
- Imputation assuming MAR will be carried out to impute missing data based on a multivariate joint Gaussian imputation model using the Markov chain Monte Carlo (MCMC) method. A separate imputation model will be used for each treatment arm. The imputation models will include countries, surgical approach, gender, age, hemodynamic assessments at each time point from Baseline. The MCMC method will be used with multiple chains, 200 burn-in iterations, and a non-informative prior. In case of non-convergence or non-estimability issues, a single model will be considered with treatment arm added as explanatory variable to the model. If case of further non-convergence the "mild" and "moderate/severe" classes will be merged.

Imputed data will consist of 200 imputed datasets. The random seed number for partial imputation with the MCMC method will be 2401, and the random seed number for the sequential regression multiple imputation will be 4181982.

The 200 imputed datasets will be used to graphically represent the trend on Leak severity (paravalvular and central) during time from baseline to 5 years by treatment arm.

12. SECONDARY COST-EFFECTIVENESS ANALYSIS

12.1. SURGICAL TIMES

12.1.1. Testing strategy

As secondary analysis, Cross clamp and Extracorporeal circulation times will be analyzed in a superiority context. A one-sided superiority test will be conducted comparing the treatment and control group at discharge in order to reject the null hypothesis that PERCEVAL is not superior to Stented valve, according to the following hypotheses:

H_0: $\mu_{PERCEVAL} \geq \mu_{CONTROL}$

H_1: $\mu_{PERCEVAL} < \mu_{CONTROL}$

where $\mu_{PERCEVAL}$ and $\mu_{CONTROL}$ are the surgical times (Cross clamp and Extracorporeal circulation times) of PERCEVAL and Control groups, respectively. Since statistical significance is needed for both time endpoints for claiming superiority on surgical times, no adjustment of multiplicity is necessary.

12.1.2. Surgical time analysis

Two independent Analyses of Covariance (ANCOVA) will be conducted to assess differences between treatments (Perceval vs Control) on:

- Cross clamp time
- Extracorporeal circulation time

after controlling for the effects of prognostic baseline/perioperative patient characteristic covariates (countries, surgical approach, concomitant procedure, gender, age).

The F-test of significance will be used to assess superiority of Perceval compared to Control. The test will be one- tailed with the probability of rejecting the null hypothesis when it is true set at $p < 0.025$.

The assumptions of normality and homogeneity of variance for performing ANCOVA will be assessed:

- Normality will be assessed using the One-Sample Kolmogorov-Smirnov test;
- Homogeneity of variance (both groups have equal error variances) will be assessed using Levene's Test for the Equality of Error Variances

As the sample size is large, small departures from normality can be detected. Because small deviations from normality do not severely affect the validity of ANCOVA tests, other statistics and plots will be examined to make a final assessment of normality:

- the skewness and kurtosis measures

- the histogram of distribution
- the Q-Q plot

As sensitivity analyses, the same ANCOVA models will be run considering also:

- previous cardiovascular procedure performed
- Medical History/Patient Risk Factors

A stepwise selection method will be used, with an Entry and Stay Significance Levels set at 0.15.

As additional sensitivity analysis, a non-parametric Kruskal-Wallis test will be performed to test the equality of medians between the treatment groups.

12.2. NORMALIZED CONSUMPTION INDEX AND REDUCED NORMALIZED CONSUMPTION INDEX

12.2.1. Definitions

According to the CIP (Appendix 3), the Consumption Index (CI) will be define as the sum of all resources consumed, weighted for the contribution of unit cost of each event to the total cost:

$$CI = \sum_{j=1}^{n} \alpha_j E_j$$

where E_j is the observed value of event j and α_j is the ratio between unit cost of event j and the total cost. The Normalized Consumption Index will be define as a transformation of CI such that it ranges between 0 and 1. The simplest definition is

$$NCI = \frac{CI}{1 + CI}$$

The contributions of each event on the total cost (α_j) are reported in Table 5 and are calculated using result of 4 European countries (Italy, France, Germany, UK).

Table 5 : Contributions of each event on the total cost

Event (within discharge)	a_j	a^*_j
OR time (hours – only first attempt)	0.1114	0.1177
Re-operation up to discharge (yes/no) [valve intervention within discharge]	0.2251	
ICU LOS (days)	0.0798	0.0843
Ward LOS (days)	0.0254	0.0268
Transfusion (units during implant and ICU)	0.0093	
Sepsis (yes/no) [from AE]	0.2544	
Was the subject under renal replacement (RRT) therapy in ICU? (yes/no)	0.0092	

Any Permanent Pacemaker Implant	0.2006	
Stroke (yes/no) [from AE]	0.2182	

Since OR time and hospital cost represent more than 95% of total cost for all countries, a reduced Normalized Consumption Index (rNCI) will be define as the previous index but using only OR time, ICU and ward length of stay ($\alpha *_j$).

Descriptive analyses for NCI, rNCI, and their components will be reported overall, by treatment and stratified by country.

12.2.2. **Testing strategy and analysis**

As co-secondary analysis, Normalized consumption index and Reduced normalized consumption index will be analyzed in a superiority context. A one-sided superiority test will be conducted comparing the treatment and control group at discharge in order to reject the null hypothesis that PERCEVAL is not superior to Stented valve, according to the following hypotheses:

H_0: $\mu_{PERCEVAL} \geq \mu_{CONTROL}$

H_1: $\mu_{PERCEVAL} < \mu_{CONTROL}$

where $\mu_{PERCEVAL}$ and $\mu_{CONTROL}$ are the Normalized and Reduced normalized consumption indexes of PERCEVAL and Control groups, respectively. Since statistical significance is needed for both time endpoints for claiming superiority on consumption index, no adjustment of multiplicity is necessary.

If less than 5% of data for consumption indexes are missing, the analyses will be based on the complete cases and the comparison will be done by means of Wilcoxon-Mann-Whitney tests.

Otherwise, a multiple imputation process will be applied. In case of non-normality, data will be transformed toward normality before imputation using a logarithmic transformation. Treatment groups will be compared using Multiple Imputation (MI) approach and an ANCOVA model. Mean values of NCI and rNCI will be analyzed based on observed and imputed data. Multiple imputation will be performed under the assumption of missing-at-random (MAR) and will be implemented in using regression multiple imputation. The regression model will include as explanatory variables the treatment, surgical approach and the presence of concomitant procedures. No rounding or range restrictions will be applied. Imputed data will consist of 300 imputed datasets. The random seed number for the regression multiple imputation will be 9877. Each of the 300 imputed datasets will be analyzed by means of an ANCOVA analysis using the observed/imputed values as the dependent variable. The model will include as fixed effect categorical factors for treatment groups. Results from analysis of each imputed dataset, i.e., LSM treatment differences and their standard errors, will be combined using Rubin's imputation rules to produce a pooled LSM estimate of treatment difference, its 95% confidence interval, and a pooled p-value for the test of null hypothesis of no treatment effect.

In case of imputation, sensitivity analyses will be performed using complete cases and Wilcoxon-Mann-Whitney tests.

12.3. HEALTH STATUS (EQ-5D-5L)

The number and proportion of subjects of each level (1 - no problems, 2 - slight problems, 3 - moderate problems, 4 - severe problems, 5 - extreme problems) for each dimension (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) will be reported. The change in numbers and proportions from baseline to post-implant FU will be calculated9. The tables will be shown by treatment arm.

The Paretian classification9 will be applied at each available post-implant FU according to the following rules:

- A health profile is better than another if it is better in at least one dimension, and is no worse in any other dimension.
- A health profile is worse than another if it is worse in at least one dimension, and is no better in any other dimension.

Using this principle to compare a patient's EQ-5D health states between any two time periods, there are only 4 possibilities:

1. Their health state is better
2. Their heath state is worse
3. Their health state is exactly the same
4. The changes in health are 'mixed': better on one dimension, but worse on another.

According to the above classification, changes from baseline in health status by treatment arm will be reported. Additionally, changes from baseline will be displayed as pie charts (by treatment and by each available post-implant assessment).

EQ-5D-5L answers to each dimension (health states) will be converted into a single index value. The index values allow the calculation of quality-adjusted life years (QALYs) that are used to inform economic evaluations of health care interventions. The idea underlying the QALY is that a year of life lived in perfect health is worth 1 QALY (1 Year of Life × 1 Utility = 1 QALY) and that a year of life lived in a state of less than this perfect health is worth less than 1. In order to determine the exact QALY value, it is requested to multiply the index value associated with a given state of health by the years lived in that state. The steps to obtain a QALY are:

1. Each dimension of the EQ-5D-5L questionnaire will be scored as follow:
    - Level 1 will be coded as '1'
    - Level 2 will be coded as '2'
    - Level 3 will be coded as '3'
    - Level 4 will be coded as '4'
    - Level 5 will be coded as '5'
    - Missing or ambiguous values (e.g. 2 boxes are ticked for a single dimension) will be coded as '9'.

All five responses are then combined for a five-digit health state. For each five-digit health state there is a correspondent health related quality of life or score. The example in Figure 2 identifies the state '12345'.

2. Each individual response to EQ-5D-5L will be converted into a single utility value by applying value sets reported in Appendix I (provided by the EuroQol website). UK coefficients will be used as primary analysis (they represent the standard). Since Germany and France will represent the top enrolling countries, the Germany-based and France-based scoring system will be used as sensitivity analyses in order to assess the robustness of results changing the country value sets. States including '9' will not be set as missing.

   For example, the state '12345' would be converted in an utility value of 0.063 using the UK value set.

3. The utility values will be calculated for each patient at each available time point. An averaged index value will be calculated for each time interval (i.e., Baseline – M3, M3 – Y1) as:

   $$QALY_{Vi \rightarrow Vj} = k*(utility_{Vi} + utility_{Vj})/2$$

   where $QALY_{Vi \rightarrow Vj}$ is the averaged quality of life between Vi and Vj, $utility_{Vi}$ and $utility_{Vj}$ are the utility values at visit Vi and Vj respectively, and k is a weight representing the portion of year that Vi-Vj covers:

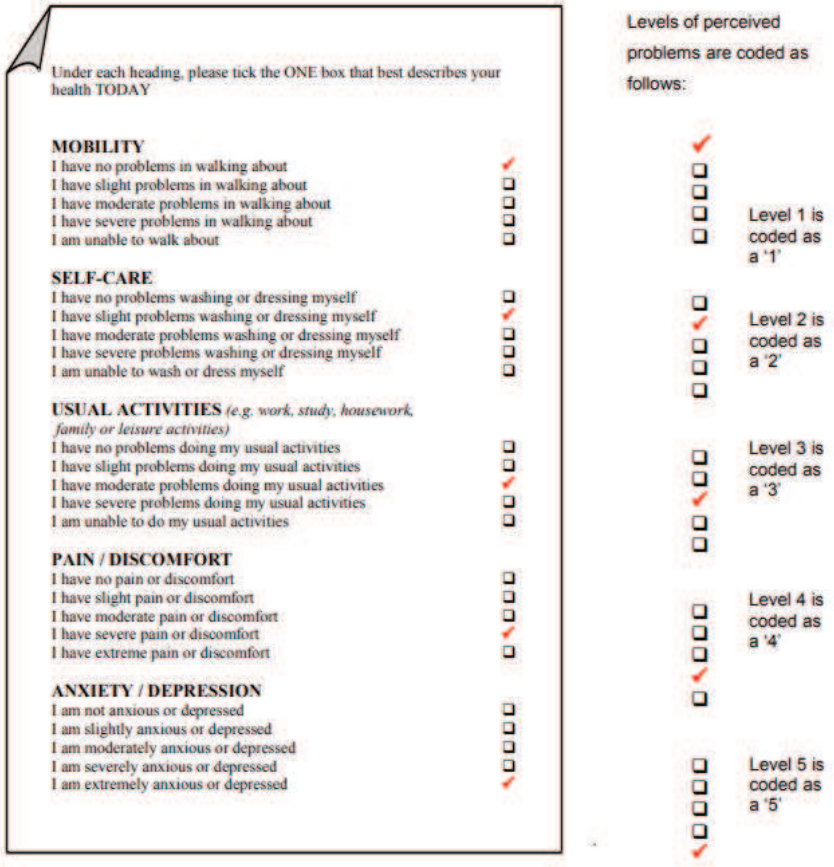
   k= (date of EQ5D at Vj – date of EQ5D at Vi)/365.25

   In order to use the $utility_{Vi}$ and $utility_{Vj}$ data, the following criteria should be met:

   - For M3: (date of EQ5D at M3 – date of EQ5D at Baseline)/365.25 is lower or equal to 0.3;
   - For Y1: (date of EQ5D at Y1 – date of EQ5D at Baseline)/365.25 is between 0.9 and 1.1.

   The quality of life between baseline to a specific time point Vi is given by the sum of all time-interval QALY from baseline to Vi. No discount rate will be applied, since evaluation is performed at 1 year.

Figure 2: Example of EQ-5D-5L form



Descriptive statistics of the change from baseline of the QALY and EQ-5D index values (continuous) at each FU assessment will be shown as described in section 5.2.

Linear mixed model for repeated measures (MMRM) will be applied to compare EQ-5D index values between the two groups. EQ-5D index values will be adjusted by baseline index value, time point, and treatment-by-time interaction. Patient will be also added to the model as a random variable. Estimates of the least square means for EQ-5D index values and 95 % confidence intervals (CIs) will be calculated by each visit and group.

Descriptive statistics of the change from baseline of the EQ VAS data (continuous) at each FU assessment will be shown as described in section 5.2.

Mann-Whitney tests will be applied to compare treatments in terms of EQ-5D index values and EQ VAS data.

13. SECONDARY SAFETY ANALYSIS

Safety data will be summarized descriptively by treatment group.

13.1. ANALYSIS OF ADVERSE EVENTS

A treatment emergent adverse event (TEAE) is defined as any AE that based on start date occurs after implant as defined in Section 5.4. If a non-TEAE worsens after implant, it has to be reported as non-TEAE until worsening and as TEAE afterwards.

A pre-treatment non-TEAEs is an AE starting in the pre-treatment period, a post-treatment non-TEAEs is an AE starting in the post-treatment period as defined in Section 5.4.

If there are partial dates or times, an AE will be considered treatment emergent unless the information available will clearly exclude it. Further details can be found in Section 18.10.

13.1.1. Adverse Events overviews

The following overview tables will be generated by treatment group and overall.

- Summary of the number and percentage of subjects with at least 1
  - Serious TEAE
  - Related serious TEAEs
  - TEAE leading to discontinuation from the study
  - Deaths

  The percentage denominator will be the number of subjects.
- Summary of the number and percentage of TEAEs for
  - TEAE
  - Serious TEAE
  - Related serious TEAEs
  - TEAE leading to discontinuation from the study

  The percentage denominator will be the total number of TEAEs.

13.1.2. Adverse Events Incidences

The incidence of AE is defined as the number of subjects with occurrence of this AE during the period of interest. The incidence rate (CIR for crude incidence rate) of an AE is defined as the number of subjects with occurrence of this AE during the period of interest divided by the total number of subjects n in the respective group (e.g., treatment group).

The incidence, incidence rate, the number of events and the percentage of events (related to the total number of events) will be summarized by AE term over the entire study post-implant period (from Day 1) and during the early and late phases (section 18.12). For the late events, the incidence rate will be calculated on the number of late patients only.

In addition, safety summaries will be done according to the following plan:

- Intraprocedural and periprocedural (within 72 hours) serious adverse events regardless of relationship with the device.

- All valve and procedure relevant (according to CEC adjudication) serious adverse event composite endpoints and their components as specified in VARC-2 guidelines (and reported in section 18.12):
    - Early safety at 30 days
    - Clinical efficacy after 30 days
    - Time related valve safety annually up to 5 years
- Serious device related (according to CEC adjudication) adverse events up to 5 years (from day 1 up to Day 1826)
- Postoperative site-reported unexpected serious adverse device effect (USADE)
- Site-reported relationship to procedure, device, drug/medication, or other Protocol Required Assessment (section 18.12)
- All late serious adverse events will be analyzed according to a constant hazard model: the linearized rates will be calculated as:
    - total number of late events divided by the total time under evaluation after 30 days (late patient-years).
    - number of patients with at least one late event divided by the total time under evaluation after 30 days (late patient-years).
- The incidence, incidence rate, the number of events and the percentage of Pacemaker (PM) implantation according to:
    - time of occurrence (overall, early, late and up to 1 year), and
    - relationship to device and procedure.

In addition, time-to-event methods (section 5.2) will be used to summarize the time to onset of the following events:

- deaths (overall, cardiovascular, and non-cardiovascular),
- explants and
- all main cardiovascular SAEs:
    - Bleeding,
    - SVD,
    - endocarditis,
    - valve thrombosis,
    - thromboembolic events,
    - stokes,
    - NSVD,

- central leak,
- perivalvular leak,
- thrombocytopenia,
- hemolysis,
- arrhythmias and conduction disorders,
- myocardial infarction, and
- hospitalizations (overall , cardiovascular hospitalization, valve-related hospitalizations);

- First PM implant.

For events leading to hospitalizations, descriptive analysis of the duration will be provided.

The following specific listings will be produced for all enrolled subjects:

- Deaths (the listing will present flags indicating the relevant study phase).
- Serious AEs other than death
- MACCE

13.2. ANALYSIS OF LABORATORY PARAMETERS

Summary statistics of blood laboratory data will be provided along with summary of the change from preoperative assessment, calculated for continuous variables as difference of values (section 18.2).

The percent of subjects with results within the lab normal values at each time interval will be calculated (normal reference limits provided by the sites). In addition, shift tables of values within normal ranges from preoperative to annual post-operative FU assessment will be generated by treatment group.

13.3. PLATELET COUNT AND CKMB DETERMINATION

Platelet count and CKMB/troponin will be analyzed by means of summary statistics (section 5.2) and box plots at preoperative and at each determination, overall and for each treatment arm.

13.4. PACEMAKER EVALUATION

For all patients who will have been implanted with a permanent pulse generator, all ECG Core Lab parameters will be descriptively analyzed by visit and treatment arm.

14. SUBGROUP ANALYSIS

The following analyses will be repeated for each subgroup (see section 5.3) using the same analysis population planned for each analysis:

- Subject disposition (section 0)

- Subject demographics (section 8.1), Implant characteristics (section 8.2), Valve characteristics (section 8.3), CT Scan (section 8.4), other baseline characteristics (section 8.6), and Healthcare Economics (section 8.8).
- Primary Endpoint (section 10).
- Forest plot showing the hazard ratio and 95% confidence intervals associated with variables considered in the multivariate analyses (treatment arm and all variables listed in section 5.3) with time to the primary endpoint (MACCE) as the dependent variable.
- Adverse Events (section 13.1)

TFLs will report per each subgroup level the overall and by treatment results. Analysis by subgroup will be performed if each subset has at least 15% of the allocated subjects.

15. INTERIM ANALYSIS

Two interim analyses will be performed to select the sample size. These interims will be triggered when the number of subjects enrolled will be approximately 900 and 1050 (interim analyses cut-off dates). The study will stop accrual for expected success if the posterior probability of non-inferiority in the PP population is sufficiently high. The threshold varies by interim: accrual will stop if the posterior probability exceeds

- 0.996 at N=900, or
- 0.993 at N=1050.

The analysis will be conducted by the external Independent Statistical Unit who will communicate to the Sponsor the result of the interim analysis and the consequent decision on the continuation of enrollment (i.e. “Stop enrolment for predicted success”, “Continue enrolment”).

As the interim analyses will only determine whether accrual needs to be stopped, then the interim analyses decisions have no impact on the subsequent study conduct. Therefore subject follow up will continue up to 5 years and the primary analysis will take place at the 12-month primary analysis.

At the time of each interim analysis, there will be subjects eligible for Per Protocol analysis population whose one-year outcome is unknown (e.g. ongoing subjects not having yet reached the one-year visit). For this reason at the interim analysis an imputation model will be used for their one-year outcome according to these rules:

- Any subjects that are missing the one-year outcome, but are known to have a MACCE event at an intermediate time point, will have their one-year outcome imputed as “not MACCE-free”.
- Any subjects that are missing the one-year outcome but have not had a MACCE event at an intermediate time point (3, 6, and 9 month intervals) will have their one-year outcome imputed, using multiple imputation. The probability that a subject remains MACCE-free at one year, conditional on being MACCE-free as of their last known intermediate time point (3, 6, and 9-month intervals), follows a Beta-Binomial distribution. A non-informative Beta(1,1) priors, separately within each arm is going to be applied. Additional details are provided in the CIP Appendix 8.

16. ABBREVIATIONS

Table 6 : List of abbreviations and acronyms

Acronym	*Verbatim*
AE	Adverse Event
ANCOVA	Analysis of Covariance
AVR	Aortic Valve Replacement
CEC	Clinical Event Committee
CI	Consumption Index
CIP	Clinical Investigation Plan
CM	Concomitant Medication
CT-Scan	Computerized Tomography Scan
DSMB	Data Safety Monitoring Board
DVI	Doppler Velocity Index
EC	Ethics Committee
ECG	Electrocardiogram
eCRF	electronic Case Report Form
EDC	Electronic Data Capture
ENR	Enrolled population
EOA	Effective Orifice Area
EOAi	Effective Orifice Area index
EQ-5D-5L	European Quality of Life - 5 Dimensions – 5 Levels
ICH	International Conference on Harmonization
ICU	Intensive Care Unit
ITT	Intention To Treat
LV	Left Ventricle
LVEF	Left Ventricle Ejection Fraction
LVOT	Left Ventricle Outflow Tract
M1	1 month assessment (EQ5D)
M3	1 to 3 months FU visit
MACCE	Major Cerebral And Cardiovascular Event
MI	Multiple Imputation

Acronym	*Verbatim*
MICS	Minimal Invasive Cardiac Surgery
mITT	modified Intention To Treat population
MMRM	Mixed Model for Repeated Measures
MPG	Mean Pressure Gradient
mRS	modified Rankin Scale
NCI	Normalized Consumption Index
NIHSS	NIH Stroke Scale/Score
NYHA	New York Heart Association
PM	Pace Maker
PP	Per Protocol population
PPG	Peak Pressure Gradient
PPI	Permament Pacemaker Implant
QALY	Quality Adjusted Life Years
QoL	Quality Of life Questionnaire
SAE	Serious Adverse Event
SAF	Safety population
SAP	Statistical Analysis Plan
STS	Society of Thoracic Surgeons
SVD	Structural Valve Deterioration
TAVI	Transcatheter Aortic Valve Implantation
TEAE	Treatment Emergent Adverse Event
TFL	Table, Figures and Listings
TTE	Trans Thoracic Echocardiogram
USADE	Unanticipated Serious Adverse Device Effect
VARC	Valve Academic Research Consortium
Y1 (Y2,Y3,Y4,Y5)	1 (2,3,4,5) year(s) FU visit

17. REFERENCES

1. Perceval Sutureless Implant Vs Standard Aortic Valve Replacement (TPS003) CIP, version 3.0 dated 27 February 2018

2. TPS003 Patient Casebook, version MSC004 dated 12 July 2018

3. Muneretto, C., et al., *Treating the patients in the 'grey-zone' with aortic valve disease: a comparison among conventional surgery, sutureless valves and transcatheter aortic valve replacement.* Interact Cardiovasc Thorac Surg, 2014.

4. Pollari, F., et al., Better short-term outcome by using sutureless valves: a propensity-matched score analysis. Ann Thorac Surg, 2014. 98(2): p. 611-6; discussion 616-7.

5. Dalén, M., et al., Aortic valve replacement through full sternotomy with a stented bioprosthesis versus minimally invasive sternotomy with a sutureless bioprosthesis. European Journal of Cardio-Thoracic Surgery, 2015.

6. Konig, K.C., et al., Sutureless Perceval aortic valve in comparison with the stented Carpentier-Edwards Perimount aortic valve. J Heart Valve Dis, 2014. 23(2): p. 253-8.

7. Mohr, F.W., et al., The German Aortic Valve Registry: 1-year results from 13 680 patients with aortic valve diseasedagger. Eur J Cardiothorac Surg, 2014. 46(5): p. 808-16.

8. Kalbfleisch, J.D.a.P., R. L. , The Statistical Analysis of Failure Time Data, ed. J.W. Sons. 1980, New York.

9. Devlin N, Parkin D, Browne J. (2010). Using the EQ-5D as a performance measurement tool in the NHS. *Health Economics* 19(8):886-905.

18. ANALYSIS SPECIFICATIONS

The purpose of this section is to give technical details for the implementation of the SAP.

18.1. PERCENTAGES AND DECIMAL PLACES

If not otherwise specified, the following rules are applied:

- Percentages are presented to 1 decimal point.
- Percentages equal to 0 or 100 are presented as such without a decimal point.
- For descriptive summary statistics, the same number of decimal places as in the raw data are presented when reporting minimum and maximum values, 1 more decimal place when reporting mean, median, quartiles and confidence interval (CI) and standard deviation (SD).
- P-values are presented to 3 decimal points. P-values < 0.001 will be reported as such.
- Ratios are presented to 3 decimal points.

The above described displaying rules must not be changed (e.g., rounding) for the CSR text and are used 1:1 in the body report as well.

18.2. PRESENTATION OF DIFFERENCES AND CHANGES

If not differently indicated:

- for differences between active and comparator, active will constitute the minuend and comparator the subtrahend. For differences between active comparator and placebo the active comparator will constitute the minuend and placebo the subtrahend;
- for changes from baseline the baseline value will constitute the minuend and the later value the subtrahend.

18.3. PRESENTATION OF UNITS

If applicable, parameters will be displayed together with the used unit of measurement. The unit of measurement is enclosed in square brackets ([ ]). This applies to both tables and listings.

18.4. PRESENTATION OF DATES

Where applicable (e.g., in listings), dates will be displayed in ISO8601 format (example: 2014-09-29T12:16, see CDISC 2013). In case of incomplete dates, the imputed one will be listed and a flag will be added to identify what have been imputed:

- D means that the day is imputed;
- M means that the day and month are imputed;
- Y means that the entire date is imputed.

18.5. HANDLING OF MISSING VALUES

At each time point/visit, all subjects still in the study are reported. Missing values will be taken into account as missing in the analysis. The number of observed values and the number of missing values must sum up to the number of subjects in the study at the respective time point/visit.

18.6. VISIT WINDOWS

The performed visits have an analysis window defined in the CIP. In order to collect all of the information provided in the CRF, the performed visits could be recalculated in real visits with larger analysis window as defined in Table 7.

Table 7 : Visit window

Visit	*Targeted study day*
1 Month assessment	Day 30 ± 7 days
1 to 3 Months visit	between Day 30 and Day 90
1 Year visit	1 year (Day 365) ± 30 days
2 Years visit	2 years (Day 731) ± 30 days
3 Years visit	3 years (Day 1096) ± 30 days
4 Years visit	4 years (Day 1461) ± 30 days
5 Years visit	5 years (Day 1826) ± 30 days

18.7. CONVERSION OF TIME INTERVALS

If a time interval was calculated in minutes, hours or days and needs to be converted into months or years the following conversion factors will be used:

- 1 month = 30.4375 days
- 1 year = 365.25 days

18.8. MANDATORY TABLES WITHOUT DATA

Recommended tables must be created. If no subject qualifies for the table, the header will be created and the table itself will be replaced by “No subject in this category”.

18.9. STORING OF IMPUTATION DATA

The result of all imputation strategies for final analyses (e.g., incomplete start dates of AE), combination of observations [(e.g., combination of consecutive AEs)] and new derived information (e.g., treatment-emergent flag) must be stored in ADaM data set.

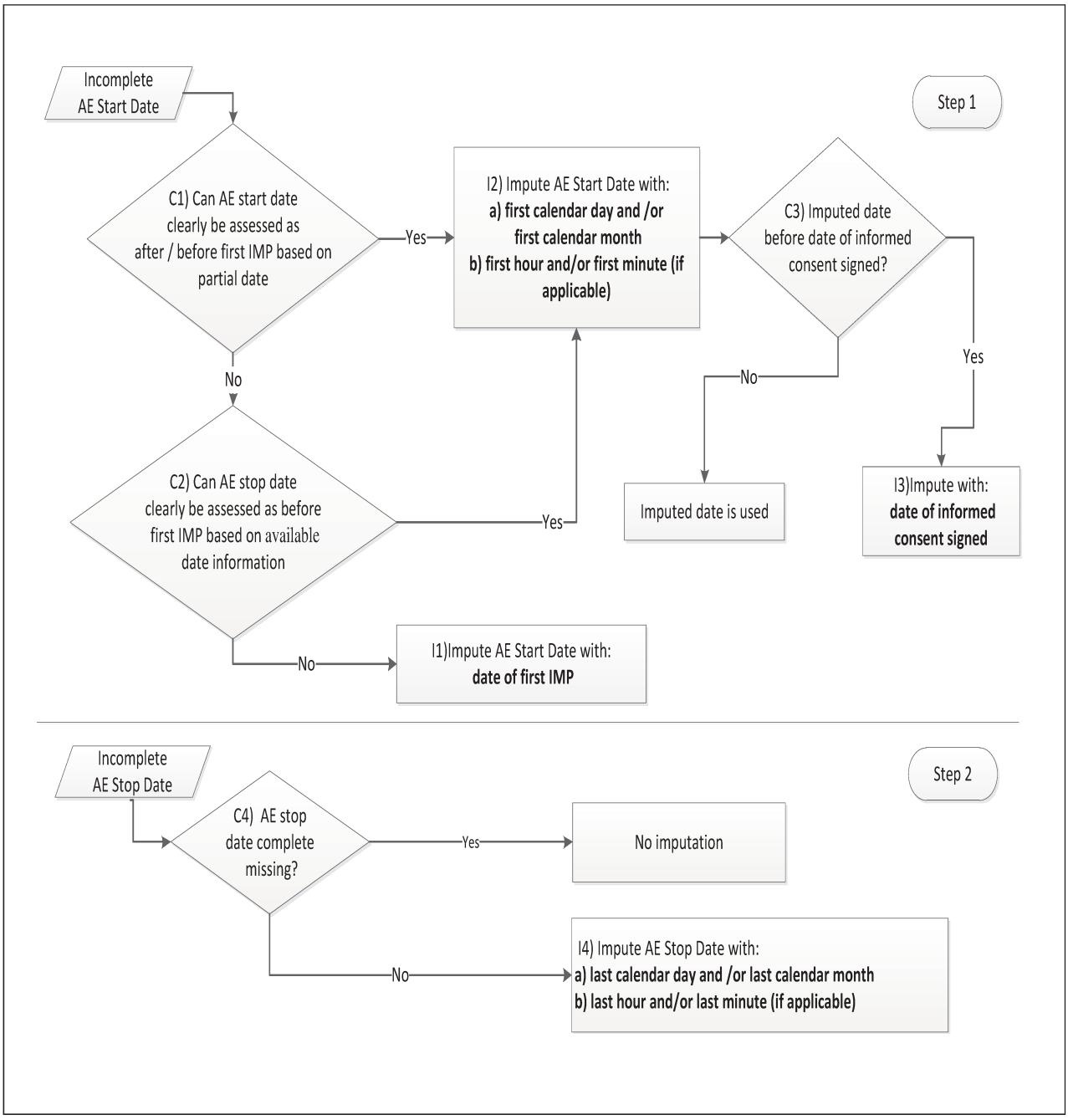
18.10. HANDLING OF MISSING DATES

The term missing date refers to a completely missing date or to an incomplete date where parts are not available. If the number of AE with missing date information does not exceed 5% the following imputation strategy is applied.

Missing start and stop date will be imputed conservatively, i.e., missing values will be imputed in such a way that the duration of the AE is considered with the longest possible duration and such that, whenever the AE may potentially start after implantation, the AE will be handled as a TEAE. Further information are reported in Section 18.12 (AE start date) and Figure 3 (AE stop date).

A replacement of missing year for AE start information is not foreseen. If needed, this will be considered on a case-by-case decision which must be documented together with the documentation of ADaM data sets.

Figure 3: Graphical overview about the imputation strategy



I1-I4: imputation steps
C1-C4: checkpoints

18.10.1. **Assessment of TEAEs**

The assessment whether an AE is a TEAE will be done after replacement of missing date information. As the on-treatment-period is defined without time information the assessment of TEAE will be based on the date part of the AEs only.

18.10.2. **List of deaths**

Deaths will be identified by SAE defined as “fatal” or “leading to death” (SAE criteria).

18.10.3. **Time to onset of AE**

Time to onset of AE will be calculated based on date of implant and on the imputed value for AE start date.

18.10.4. **Duration of AE**

Duration of AE will be calculated based on the imputed values for AE start date and stop date.

If duration of AE could not be calculated due to unknown date information the following assessment to categories will be used:

- If the AE is marked as “ongoing” in the e-CRF the duration will be categorized as “ongoing”
- Otherwise the duration category will be set to “missing”.

18.11. TITLES AND FOOTNOTES

Footnotes should include:

- Analysis cut-off date
- The name and version of the coding system for TFLs presenting coded items.

18.12. DATA DERIVATIONS

Table 8 : Data derivation rules

Derived variable	*Derivation algorithm*
Consented subject	Yes if filled date of consent No, otherwise.
Enrolled subject	Yes if filled date of consent AND patient is not a failure to implant (termination reason= FAILURE TO MEET RANDOMIZATION CRITERIA) No, otherwise.
ITT subject	*Yes if Randomization date not empty* No, otherwise.
mITT subject	*Yes if ITT and the valve is a study valve (as listed in CRF). In case of "other" values, they have to be checked one by one with PM.* No, otherwise.
Adherence to Randomization allocation	*Yes if "Was the randomized valve implanted?" at first attempt=Yes* No, otherwise.
SAF subject	*Yes if Randomization andImplant dates not empty* No, otherwise.
PP subject	*Yes if* - *Randomization andImplant dates not empty* - Adherent to randomization - no major protocol violation (as reported in CRF field). No, otherwise.

Derived variable	***Derivation algorithm***
Operating room suite time (min)	*Implant, Surgical time:* Finish time – Start time
Skin-to-skin time (min)	*Implant, Surgical time:* Finish time – Start time
Extracorporeal circulation time (min)	*Implant, Surgical time:* Finish time – Start time
Cross clamp time (min)	*Implant, Surgical time:* Finish time – Start time
ICU stay (minutes)	*Intensive Care Unit:* [Date/Time of ICU Discharge] - [Date/Time of ICU Admission] *If time of ICU discharge is missing it will be set at 23:59* *If time of ICU Admission is missing it will be set at 00:00*
ICU stay (days)	*ICU stay (minutes) / 1440*
Time to SAE	*Serious Adverse Event (SAE):* SAE start date – Surgery date +1
Early, Late SAE	EARLY if Time to SAE<=30 days LATE if Time to SAE > 30 days

Derived variable	***Derivation algorithm***
Early safety (at 30 days)	*Serious Adverse Event (SAE):* *YES if at least one* (Time to SAE <=30 days) *per CEC:* - All-cause mortality - All stroke disabling and nondisabling - Life-threatening bleeding - Acute kidney injury—Stage 2 or 3 (including renal replacement therapy) - Coronary artery obstruction requiring intervention - Major vascular complication - Valve-related dysfunction requiring repeat procedure (BAV, TAVI, or SAVR): SVD and Aortic regurgitation requiring re-intevention *NO, otherwise*
Clinical efficacy (after 30 days)	*Serious Adverse Event (SAE):* *YES if at least one (Time to SAE >30 days) per CEC or Echo core lab:* - All-cause mortality - All stroke (disabling and nondisabling) - Requiring hospitalizations for valve-related symptoms or worsening congestive heart failurey. - NYHA class III or IV (date of assessment >30 days after valve implant) - Valve-related dysfunction: i. (mean aortic valve gradient ≥20 mm Hg, EOA ≤0.9–1.1 cm2 and/or DVI<0. 35 m/s, AND/OR ii. moderate or severe prosthetic valve regurgitation) *NO, otherwise*

Derived variable	*Derivation algorithm*
Time-related valve safety	*Serious Adverse Event (SAE):* *YES if at least one (Time to SAE >30 days) per CEC or echo core lab:* - Structural valve deterioration (both points i. and ii. have to be met): i. Valve-related dysfunction: ▪ mean aortic valve gradient ≥20 mm Hg, EOA ≤0.9(if BSA<1.6 cm2) or 1.1(if BSA≥1.6 cm2) cm2 and/or DVI<0.35 m/s, AND/OR ▪ moderate or severe prosthetic valve regurgitation) ii. Requiring repeat procedure (TAVI or SAVR) - Prosthetic valve endocarditis - Prosthetic valve thrombosis - Thrombo-embolic event, non-cerebral , - TIA - STROKE - VARC bleeding (BARC type 5, BARC type 3A, 3B and 3C, BARC type 2), unless clearly unrelated to valve therapy (e.g., trauma) : Device and/or procedure related *NO, otherwise*
MACCE (Yes/No)	*Serious Adverse Event (SAE), Valve re-intervention:* *YES if at least one per CEC:* - Death - Myocardial infarction - Stroke - Reintervention *NO, otherwise*

Derived variable	***Derivation algorithm***
MACCE @ 1 year (Yes/No)	*Serious Adverse Event (SAE), Valve re-intervention:* *YES if at least one per CEC:* - Death (Time to SAE <=365 days) - Myocardial infarction (Time to SAE <=365 days) per CEC - Stroke (Time to SAE <=365 days) per CEC - Reintervention (Time to SAE <=365 days) *NO, otherwise*
MACCE @ n year (Yes/No)	*CEC adjudication of Serious Adverse Event (SAE), Valve re-intervention:* *YES if at least one per CEC:* - Death (Time to SAE <=n*365.25 days) - Myocardial infarction (Time to SAE <= n*365.25 days) per CEC - Stroke (Time to SAE <= n*365.25 days) per CEC - Reintervention (Time to SAE <= n*365.25 days) *NO, otherwise*
PM implant	*SAE FORM – TREATMENT* *Yes if:* - Was a pulse generator implanted? = YES - If yes, specify = Pacemaker *NOTE: PM implant will be classified as intraoperative, early or late according to the date of PM implant (not the AE start date)*

Derived variable	***Derivation algorithm***
Last exposure date	*Study Termination, Serious Adverse Event (SAE):* 1. For explant of investigational device: date of explant 2. For dead subjects: date of death 3. For alive and not explanted subjects: - Date of termination - If there is no date of termination, the last available FU date (including discharge) or the last SAE date (occurrence date), whichever occurs later. - If there is no follow-up and no adverse events : date of surgery or date of randomization, whichever occurs later.
Last study date	*Study Termination, Serious Adverse Event (SAE):* 1. For dead subjects: date of death 2. For alive subjects: - Date of termination - If there is no date of termination, the last available FU date (including discharge) or the last SAE date (occurrence date), whichever occurs later. If there is no follow-up and no adverse events : date of surgery or date of randomization, whichever occurs later.
Study length (days)	*Study termination, Subject Information:* Last study date – Enroll Date+1
Extent of exposure (days)	Last exposure date-surgery date+1
Late patients	Subjects with more than 30 FU days of exposure
Patient-years	Cumulative Study length (days)/365.25

Derived variable	*Derivation algorithm*
Early patient-years	Cumulative Study length (days) up to 30 days/365.25
Late patient-years	Cumulative Study length (days) from 31st day/365.25. Subjects with less than 31 FU days do not contribute to the late patient-years
SAE starting date (missing)	In case of missing SAE starting date, the following rule will be applied: 1. If missing day: 1st day of the month 2. If missing day and month: 01-JAN 3. If totally missing: Date of surgery SAE starting date cannot be before the surgery date. In that case, SAE starting date will be set as date of surgery.
Deaths	Death is when there is a SAE: - With outcome='FATAL' OR - Leading to death (SAE criteria) The date of death is retrieved from the SAE form (Date of death).
Leak severity	Leak will be classified in: 1. none/trace: if "Is there a Paravalvular Leak?" ("Is there a Central Leak?")='No' or severity=' Trace (<1/4+)' 2. mild: if severity=' Mild (1/4+)' 3. moderate/severe: if severity=' Moderate (2/4+)' or severity=' Severe (>= 3/4+)'

Derived variable	*Derivation algorithm*
Proportion of patient event-free at one year	For each MACCE component: Numerator: #patients who experienced the specific MACCE component within DAY 365 Denominator: #patients who experienced the specific MACCE component within DAY 365 + # patients without the specific event at DAY 365
Proportion of patient MACCE-free at each year	Numerator: #patients who experienced MACCE within DAY 365 (730, 1095, 1460, 1825) Denominator: #patients who experienced MACCE within DAY 365 (730, 1095, 1460, 1825) + # patients without MACCE at DAY 365 (730, 1095, 1460, 1825)
MPG Gradient class	Aortic Mean Pressure Gradient (MPG) class: • MPG ≥ 20mmHG • MPG < 20mmHG
Hospitalization time [days]	Discharge – Hospital Admission +1
Ward time [days]	Hospitalization time – ICU time
EQ-VAS	Use as first source of values the "EQ5D02-Your Health Today" variable. If it's missing, use the "EQ VAS Score" to replace the missing value.
Aortic Mean Pressure Gradient (MPG) with LVOT adjustment	Aortic Mean Pressure Gradient - LVOT Mean Pressure Gradient
Peak Mean Pressure Gradient (MPG) with LVOT adjustment	Aortic Peak Pressure Gradient - LVOT Peak Pressure Gradient