

**Confidential and Proprietary**

Nothing herein is to be disclosed without the expressed written consent of  
Abbott Vascular

Protocol 10-389

**EXCEL Clinical Trial**

*Evaluation of XIENCE versus Coronary Artery Bypass Surgery  
for Effectiveness of Left Main Revascularization*

**Statistical Analysis Plan  
(Part I: Methodology)  
Version 6.0  
January, 26 2016**

**Melissa Nichols  
Aurora Breazna**

## TABLE OF CONTENTS

<b>1</b>	<b>ACRONYMS AND ABBREVIATIONS.....</b>	<b>4</b>
<b>2</b>	<b>SYNOPSIS OF STUDY DESIGN AND PROCEDURES .....</b>	<b>6</b>
2.1	PURPOSE OF THE STATISTICAL ANALYSIS PLAN .....	6
2.2	OBJECTIVE OF THE STUDY .....	6
2.3	DESIGN OF THE STUDY .....	6
2.3.1	Primary Endpoint.....	9
2.3.2	Major Secondary Endpoints.....	10
2.3.3	Other Secondary Endpoints .....	11
2.3.4	Quality of Life.....	13
2.4	ANALYSIS POPULATIONS .....	13
2.4.1	Intent-to-Treat Population.....	13
2.4.2	Per-Protocol Population .....	13
2.4.3	As-Treated Population .....	13
2.5	SAMPLE SIZE CALCULATIONS .....	14
2.5.1	Primary Endpoint.....	14
2.5.2	Major Powered Secondary Endpoints:.....	15
2.6	JUSTIFICATION OF DELTA.....	16
2.7	STUDY SUCCESS.....	16
2.8	RANDOMIZATION AND BLINDING.....	17
<b>3</b>	<b>ANALYSIS CONSIDERATIONS.....</b>	<b>18</b>
3.1	GENERAL STATISTICAL METHODS.....	18
3.2	ANALYSIS OF THE PRIMARY ENDPOINT.....	18
3.3	ANALYSIS OF THE POWERED SECONDARY ENDPOINTS .....	19
3.4	ANALYSIS OF THE SECONDARY ENDPOINTS (NOT POWERED) .....	20
3.4.1	Continuous Variables.....	20
3.4.2	Categorical Variables.....	21
3.4.3	Time to Event Variables .....	22
3.4.4	Hypothesis Testing.....	23
3.5	SUBGROUPS FOR ANALYSIS .....	24
3.6	EXPLORATORY ANALYSIS.....	26
3.6.1	Exploratory Composite Endpoint Analysis .....	26
3.6.1.1	Weighted Composite Endpoint.....	26
3.6.1.2	The Win Ratio.....	26
3.6.1.3	Covariate Adjusted Analysis.....	27
3.6.1.4	Other Exploratory Analyses.....	29
3.7	FOLLOW-UP TIME POINTS .....	29
3.8	HANDLING OF MISSING DATA.....	29
3.9	ASSESSMENT OF DATA POOLABILITY .....	29
3.9.1	Multiple Center Effect .....	29
3.9.2	Multiple Region Effect .....	29
3.9.3	Multiple Study Stent Effect .....	30
3.10	MULTIPLICITY ISSUES .....	30

3.11	INTERIM ANALYSIS .....	32
3.12	DOCUMENTATION AND OTHER CONSIDERATIONS .....	32
<b>4</b>	<b>VARIABLES FOR ANALYSIS.....</b>	<b>33</b>
4.1	ENROLLMENT AND SUBJECT DISPOSITION .....	33
4.2	DEMOGRAPHICS, SUBJECT CHARACTERISTICS, AND PROCEDURE INFORMATION .....	33
4.2.1	Demographics .....	33
4.2.2	Risk Factors .....	33
4.2.3	Cardiac History .....	33
4.2.4	Physical Measurements.....	34
4.2.5	Disease State/Severity.....	34
4.2.6	Laboratory Tests .....	34
4.2.7	Procedure/Device Information.....	34
4.2.7.1	Procedure/Device Information for PCI arm only.....	35
4.2.7.2	Procedure Information for CABG arm only .....	35
4.2.7.3	Device Malfunction .....	35
4.2.8	Protocol Medications .....	35
4.2.9	Angina Status.....	36
4.3	MORPHOLOGY AND QUANTITATIVE CORONARY ANGIOGRAPHY .....	36
4.3.1	SYNTAX Score .....	36
4.3.2	Morphology.....	36
4.3.3	Quantitative Coronary Angiography.....	36
4.4	SAFETY ANALYSIS .....	37
<b>5</b>	<b>CHANGES FROM PROTOCOL.....</b>	<b>38</b>
<b>6</b>	<b>REFERENCES.....</b>	<b>39</b>
<b>7</b>	<b>APPENDICES.....</b>	<b>40</b>
	APPENDIX A: PLANNED ANALYSES FOR IFU .....	40

## 1 ACRONYMS AND ABBREVIATIONS

Acronym/Abbreviation	Definition
ACEF	Age, Creatinine, and Left Ventricle Ejection Fraction
AE	Adverse Event
ARC	Academic Research Consortium
CABG	Coronary Artery Bypass Graft
CASS	Coronary Artery Surgery Study
Cm	Centimeter
CEC	Clinical Events Committee
CI	Confidence Interval
CKD	Chronic Kidney Disease
CTO	Chronic Total Occlusion
DS	Diameter Stenosis
EU	European Union
FFR	Fractional flow reserve
GPIIb/IIIa	Glycoprotein IIb/IIIa inhibitors
HR	Hazard Ratio
IFU	Instructions for Use
ITT	Intent-To-Treat
IVUS	Intravascular Ultrasound
Kg	Kilogram
LVEF	Left Ventricular Ejection Fraction
MACE	Major Adverse Cardiac Events
MI	Myocardial Infarction
m	Meter
mm	Millimeter
mmHg	Millimeter of mercury
MLA	Minimum Lumen Area
MLD	Minimum Lumen Diameter
NCSS/PASS	Number Cruncher Statistical System/Power Analysis and Sample Size software
OUS	Outside United States
PCI	Percutaneous Coronary Intervention
PP	Per-Protocol
QCA	Quantitative Coronary Angiography
RCA	Right Coronary Artery



<b>Acronym/Abbreviation</b>	<b>Definition</b>
RCT	Randomized Controlled Trial
RVD	Reference Vessel Diameter
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis Software
SE	Standard Error
STS	Surgical Thoracic Society
TIMI	Thrombolysis In Myocardial Infarction
TLR	Target Lesion Revascularization
TVR	Target Vessel Revascularization (TLR and TVR, non-target lesion)
ULMCA	Unprotected Left Main Coronary Artery
US	United States
XIENCE PRIME EECS	XIENCE PRIME Everolimus Eluting Coronary Stent
XIENCE V EECS	XIENCE V Everolimus Eluting Coronary Stent

## **2 SYNOPSIS OF STUDY DESIGN AND PROCEDURES**

### **2.1 PURPOSE OF THE STATISTICAL ANALYSIS PLAN**

This statistical analysis plan (SAP) is intended to provide a detailed and comprehensive description of the planned methodology and analysis to be used for the randomized component of Protocol 10-389, the EXCEL clinical study. This plan is based on Version 9.0 of the study protocol.

### **2.2 OBJECTIVE OF THE STUDY**

The objective of the study is to establish the safety and efficacy of the commercially approved XIENCE Family Stent System (inclusive of XIENCE PRIME, XIENCE V, XIENCE Xpedition and XIENCE PRO [for use outside the U.S. [OUS] only]) in subjects with unprotected left main coronary artery (ULMCA) disease (either isolated to the left main trunk or associated with disease in other coronary arteries) by demonstrating that compared to coronary artery bypass graft surgery, treatment of the left main stenosis ± other significant coronary lesions with the XIENCE stent will result in non-inferior or superior rates of the composite measure of all-cause death, myocardial infarction or stroke at the anticipated median follow-up of three years.

### **2.3 DESIGN OF THE STUDY**

The randomized control portion of this study (that will be referred to as RCT) is a prospective, unblinded, randomized multicenter trial of approximately 1,900 subjects enrolled at approximately 165 US and OUS centers. Following diagnostic angiography demonstrating significant ULMCA disease and consensus of the local Heart Team (qualified participating interventional cardiologist and cardiac surgeon) that the subject meets the study entry criteria, subjects will be consented and randomized 1:1 to: a) PCI using XIENCE (N=950), or b) CABG (N=950). Follow-up for all randomized subjects will continue for five years with a potential for additional follow-up to 10 years. The primary endpoint will be assessed at least 2-years after the last subject is randomized, with a median follow-up duration in all subjects of at least 3 years.

All randomized subjects will have a follow-up telephone contact or office visit at 30 days, 180 days, and 1, 2, 3, 4 and 5 years following the index procedure. In addition, to minimize bias in the assessment of the primary endpoint, at the time the last randomized subject reaches the 2-year follow-up duration ( $730 + 28 = 758$  days), an additional adverse event (AE) check will be performed in order for data up to 3 years ( $365 \text{ days} \times 3 = 1095$  days) from the date of randomization to be collected equally in both arms.

The AE check will commence the day after the 2-year follow-up duration of the last subject enrolled (i.e. 758 days), and will be completed prior to the database snapshot. At this time,

subjects with either of the following conditions should be contacted by each clinical site by phone or office visit to collect AE information:

- Subjects who have not yet completed the 3-year visit, unless the most recent follow-up was completed within 28 days of the AE check initiation; or
- Subjects who have completed a 3-year visit before day 1067 after randomization (i.e. >28 days prior to exactly 3 years [1095 days]).

For subjects who have completed any follow-up visit at  $\geq (1095 - 28 = 1067)$  days, the AE check is not required.

An additional group of approximately 1000 consecutive subjects who are not eligible for randomization or for other reasons are not randomized will be consented for the Universal Registry. All patients with left main disease without prior CABG in whom the visual estimated diameter stenosis is greater than or equal to 50% will be eligible for the enrollment into EXCEL registry. These subjects will be consented for a Universal Registry, and followed until the time of initial treatment per standard of care with either PCI, CABG or medical therapy.

Approximately 100 consecutive subjects from the Universal Registry with a  $\geq 50\%$  and  $<70\%$  visually estimated angiographic diameter stenosis who otherwise meet all enrollment criteria, but without significant ischemia by noninvasive testing consistent with significant ULMCA disease, and in whom IVUS shows a MLA  $>6.0 \text{ mm}^2$  and/or have an FFR  $>0.80$ , will not be randomized but will be analyzed separately as intermediate lesion subjects, and followed until the time of initial treatment per standard of care with either PCI, CABG or medical therapy.

For PCI the study device is:

#### **XIENCE V EECS Product Sizes**

<b>Diameter→ Length ↓</b>	<b>2.25 mm</b>	<b>2.5 mm</b>	<b>2.75 mm</b>	<b>3.0 mm</b>	<b>3.5 mm</b>	<b>4.0 mm</b>
8 mm	X	X	X	X	X	X
12 mm	X	X	X	X	X	X
15 mm	X	X	X	X	X	X
18 mm	X	X	X	X	X	X
23 mm	X	X	X	X	X	X
28 mm	X	X	X	X	X	X

#### **XIENCE PRIME EECS Product Sizes**

<b>Diameter → Length ↓</b>	<b>2.25 mm</b>	<b>2.5 mm</b>	<b>2.75 mm</b>	<b>3.0 mm</b>	<b>3.5 mm</b>	<b>4.0 mm</b>
8 mm	X	X	X	X	X	X
12 mm	X	X	X	X	X	X

<b>Diameter → Length ↓</b>	<b>2.25 mm</b>	<b>2.5 mm</b>	<b>2.75 mm</b>	<b>3.0 mm</b>	<b>3.5 mm</b>	<b>4.0 mm</b>
15 mm	X	X	X	X	X	X
18 mm	X	X	X	X	X	X
23 mm	X	X	X	X	X	X
28 mm	X	X	X	X	X	X
33 mm	n/a	X	X	X	X	X
38 mm	n/a	X	X	X	X	X

### **XIENCE Xpedition Product Sizes**

<b>Diameter → Length ↓</b>	<b>2.25 mm</b>	<b>2.5 mm</b>	<b>2.75 mm</b>	<b>3.0 mm</b>	<b>3.25 mm</b>	<b>3.5 mm</b>	<b>4.0 mm</b>
8 mm	X	X	X	X	X	X	X
12 mm	X	X	X	X	X	X	X
15 mm	X	X	X	X	X	X	X
18 mm	X	X	X	X	X	X	X
23 mm	X	X	X	X	X	X	X
28 mm	X	X	X	X	X	X	X
33 mm	n/a	X	X	X	X	X	X
38 mm	n/a	X	X	X	X	X	X

### **XIENCE PRO Product Sizes**

<b>Diameter → Length ↓</b>	<b>2.25 mm</b>	<b>2.5 mm</b>	<b>2.75 mm</b>	<b>3.0 mm</b>	<b>3.5 mm</b>	<b>4.0 mm</b>
8 mm	X	X	X	X	X	X
12 mm	X	X	X	X	X	X
15 mm	X	X	X	X	X	X
18 mm	X	X	X	X	X	X
23 mm	X	X	X	X	X	X
28 mm	X	X	X	X	X	X
33 mm	n/a	X	X	X	X	X
38 mm	n/a	X	X	X	X	X

### 2.3.1 Primary Endpoint

The primary endpoint is the composite measure of all-cause death, myocardial infarction (MI) or stroke (modified Rankin Scale (mRS)  $\geq 1$  and increase by  $\geq 1$  from baseline). measured by the Kaplan-Meier failure rate estimate at 3 years. For the primary analysis of the primary endpoint, this Kaplan-Meier rate will be defined from the date of randomization. All available data through 3-year follow-up will be used, but only events occurring up to day 1095 after randomization will be counted for the primary endpoint. For a sensitivity analysis, the Kaplan-Meier rate will be defined from the time of the index procedure date. Both of these analyses will be conducted at the time when approximately 50% of the patients have reached the 3-year follow-up AND all patients have reached the 2-year follow-up, whichever time point is latest.

For the primary and sensitivity analyses, days from the randomization date and procedure date to the first occurrence of one of the events within this composite, an early withdrawal date, or the last follow-up date will be calculated for each patient respectively. For patients who dropped out early, but had a death date available per the national death registry, this date will be used in the calculation of days (and will be counted as a primary endpoint event). If a patient does not have an event or an early withdrawal date:

- Days from randomization will be set to 1095 days (3 years) if their last follow-up date is known to occur after 3 years.
- The date of the last visit will be the date of the last phone call, office visit or medical contact.

The primary endpoint will first be tested for noninferiority with a margin of 4.2%. The critical value will be generated using the Com-Nougue<sup>1</sup> approach to estimating the Z-statistic with Greenwood's formula for estimating the standard error.

$$Z = \frac{F_{CABG}(T) - F_{PCI}(T) - \delta}{\sqrt{\text{Var}(F_{CABG}(T)) + \text{Var}(F_{PCI}(T))}}$$

Where, for each treatment group (CABG or PCI):

$F(T)$  = Kaplan Meier estimator of failure at time T

$\delta$  = non-inferiority margin

Var = variance of the estimator =  $\sigma^2$ , where  $\sigma$  is defined as:

$$\sigma = \text{standard error by Greenwood's formula} = S(T=1095) \sqrt{\sum_{j=1}^{1095} \frac{d_j}{n_j S_j}}$$

$S = 1 - F$  = KM estimate of survival at time T

$n$  = number of subjects

$d$  = number of events

$s = n - d$

The 95% CI for  $F_{CABG}(T) - F_{PCI}(T)$  is then:

$$F_{CABG}(T) - F_{PCI}(T) \pm z_{1-\alpha} \sqrt{Var(F_{CABG}(T)) + Var(F_{PCI}(T))}$$

P-value = 1 - probnorm(Z)

Superiority testing of the primary endpoint will be performed using the same Z-statistic calculation that is defined above, without the non-inferiority margin. Statistical significance will be evaluated for both analyses at a 1-sided alpha level of 0.025.

### 2.3.2 Major Secondary Endpoints

The primary analysis for these two powered secondary endpoints will be based on the analysis applied to the time from randomization estimate. **Sensitivity analyses will be performed on the time from procedure estimate.** There are two major powered secondary endpoints:

- Composite measure of all-cause mortality, MI or stroke (mRS $\geq 1$  and increase by  $\geq 1$  from baseline) presented as a Kaplan Meier rate defined from randomization to 30 days post randomization. This is the same as the primary endpoint measured at 30 days. Two sets of analyses will be performed. Days from the randomization date or procedure date to the first occurrence of the one of the events within this composite, an early withdrawal date, or the last follow-up date will be calculated for each patient respectively. For patients who dropped out early, but had a death date available per the national death registry, this date will be used in the calculation of days (and will be counted as an event). If a patient does not have an event or an early withdrawal date then days from randomization will be set to 30 days if their last follow-up date is known to occur after 30 days. These KM estimates will be put into the Z statistic defined in Section 2.3.1 using a margin of 2% to test for non-inferiority.
- Composite measure of all-cause mortality, MI, stroke (mRS $\geq 1$  and increase by  $\geq 1$  from baseline), or unplanned revascularization for ischemia occurring from the time of randomization up to study day 1095, as estimated by the Kaplan-Meier failure rate and analyzed using a Z statistic in the same way it is described under the Section 2.3.1, with a 1-sided alpha of 0.05 and a margin of 8.4%. All available data through 3-years of follow-up will be used. Days from the randomization date or the procedure date to the first occurrence of one of the events within this composite, an early withdrawal date, or the last follow-up date occurring by 3 years post randomization will be calculated for each patient. For patients

who dropped out early, but had a death date available per the national death registry, this date will be used in the calculation of days and counted as an event. If a patient does not have an event or an early withdrawal date:

- Days from randomization will be set to 1095 days (3 years) if their last follow-up date is known to occur after 3 years.
- The date of the last visit will be the date of the phone call, office visit or last medical contact.

### 2.3.3 Other Secondary Endpoints

The following additional endpoints will be analyzed:

- The endpoint of all-cause death, myocardial infarction or stroke, in-hospital, at 30 days, 180 days, and at 1, 2, 4 and 5 years.
- The endpoint of all-cause death, myocardial infarction, stroke or unplanned revascularization for ischemia in-hospital, at 30 days, 180 days, and 1, 2, 4 and 5 years.
- All cause mortality, cardiac death and non-cardiac death in-hospital, at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
- Protocol-defined MI at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
- All MI (periprocedural, spontaneous, Q-wave, and non Q-wave) at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
- MI adjudicated per Universal MI definition at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
- Protocol-defined stroke at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
- All stroke, ischemic stroke, and hemorrhagic stroke in-hospital, at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
- Disability following stroke event at 90 days± 2 weeks
- Complete revascularization at baseline procedure, anatomic and functional
- Ischemia-driven revascularization in-hospital, at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
  - All Ischemia-driven repeat revascularization procedures.
  - Ischemia-driven target lesion revascularization (TLR)
  - Ischemia-driven target vessel revascularization (TVR)
  - Ischemia-driven non target vessel revascularization (Non-TVR)

- All revascularization (ischemia driven and non-ischemia driven) at index procedure, in-hospital, at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
  - All repeat revascularization procedures
  - All TLR
  - All TVR
  - All non-TVR
- Stent thrombosis (ARC definition) symptomatic or asymptomatic – PCI arm only in-hospital, at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
- Symptomatic graft stenosis or occlusion (since this requires angiographic documentation, this endpoint will be compared to symptomatic ARC definite stent thrombosis) – CABG arm only in-hospital, at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
- Bleeding complications at 30 days, 180 days, and 1, 2, 3, 4 and 5 years.
  - Requirement for blood product transfusion
  - TIMI scale (major or minor)
  - BARC scale
- Time from randomization to procedure; time from index procedure to discharge; ICU days; time from index procedure to return to work
- Major adverse events (MAE) defined as composite of the following components. MAE will be assessed in-hospital and at 30 days only.
  - death
  - myocardial infarction
  - stroke
  - transfusion of  $\geq 2$  units of blood
  - TIMI major or minor bleeding
  - major arrhythmia
  - unplanned coronary revascularization for ischemia
  - any unplanned surgery or therapeutic radiologic procedure
  - renal failure
  - sternal wound dehiscence
  - infection requiring antibiotics for treatment
  - intubation for > 48 hours
  - post-pericardiotomy syndrome.



### **2.3.4 Quality of Life**

Quality of Life and Health Economics will be conducted during the trial. Analyses of these data will be described and conducted elsewhere.

## **2.4 ANALYSIS POPULATIONS**

Hypothesis testing for the primary endpoint and the major powered secondary endpoints will be performed based on the Intent-to-Treat (ITT), As treated and Per-protocol (PP) populations.

Primary analysis will be based on ITT. Other endpoint analyses, including the Economic Outcomes and Cost-Effectiveness Analysis will also be performed on both the ITT and PP populations. AEs collected for safety will only be summarized for the ITT population.

If there is any disagreement between the analysis results of the primary endpoint based on the three populations, potential causes for the disagreement will be investigated by careful examination of the clinical data. Information to be examined will include, but not be limited to, the number of enrolled subjects who did not meet all inclusion/exclusion criteria, the number of subjects and the justification for those who received a different treatment than that assigned by randomization, delays in receiving treatment, and the number of subjects lost to follow-up, by randomized treatment group.

### **2.4.1 Intent-to-Treat Population**

The ITT population will consist of all subjects randomized in the study, regardless of the treatment actually received. Subjects will be analyzed in the treatment group to which they were randomized, regardless of actual treatment received.

### **2.4.2 Per-Protocol Population**

The PP population consists of subjects with the following characteristics:

- received the initial treatment to which they were randomized
- assigned treatment must have been their first revascularization
- procedure performed within 4 weeks of randomization
- no violation of any inclusion-exclusion criteria

### **2.4.3 As-Treated Population**

The as-treated population consists of subjects with the following characteristics:

- subjects who received a protocol defined treatment

Subjects will be included in the treatment arm corresponding to the first study treatment actually received.

## 2.5 SAMPLE SIZE CALCULATIONS

Calculations were performed to determine the minimum sample size required for this study in order to provide approximately 80% power for demonstrating non-inferiority of the primary endpoint. The assumed true event rates used in the calculations of all powered endpoints were estimated based on the SYNTAX<sup>2</sup> trial and on results from previous studies of the XIENCE V® stent. All powered non-inferiority endpoints will be analyzed using the Com-Nougue approach<sup>1</sup>, using Kaplan-Meier estimates calculated from time from randomization.

### 2.5.1 Primary Endpoint

The primary endpoint of all cause death, MI or stroke ( $mRS \geq 1$  and increase by  $\geq 1$  from baseline) occurring up to 3 years post randomization will be evaluated using the difference in Kaplan-Meier failure rates in the intent-to-treat population. For the primary endpoint, sample size calculations for the Com-Nougue approach<sup>1</sup>, which utilizes the difference in Kaplan-Meier failure rate estimates, were derived using simulations. The Com-Nougue approach was very similar to that calculated using the asymptotic approach (difference  $\leq 0.5\%$ ). Therefore, the asymptotic test was used to approximate the power for the Com-Nougue approach using PASS 2008 (NCSS, LLC, Kaysville, Utah)<sup>3</sup>.

The hypothesis test is designed to show non-inferiority of PCI to CABG for the primary endpoint via the Z statistic, which is a normal approximation to the Binomial, with a one-sided alpha of 0.025. The null ( $H_0$ ) and alternative ( $H_A$ ) hypotheses are:

$$H_0: F_{\text{PCI-PE}}(T) - F_{\text{CABG-PE}}(T) \geq \Delta_{\text{PE}}$$

$$H_A: F_{\text{PCI-PE}}(T) - F_{\text{CABG-PE}}(T) < \Delta_{\text{PE}}$$

$F_{\text{PCI-PE}}$  and  $F_{\text{CABG-PE}}$  are the Kaplan-Meier estimates of failure rate of the primary endpoint at 3 years in the PCI and CABG arms, respectively.  $\Delta_{\text{PE}}$  is the non-inferiority margin for the primary endpoint.

The sample size calculation is based on the following assumptions:

- primary endpoint event rate is 11% in each treatment arm at 3 years (using 3-year event rates from the SYNTAX trial, the most contemporary reference dataset)
- minimum time to follow-up is 2 years
- median time to follow-up is approximately 3 years
- 8% lost to follow-up at 3 years
- non-inferiority margin  $\Delta_{\text{PE}} = 4.2\%$
- one-sided alpha = 0.025
- accrual time of 29 months.

A sample size of 1,900 subjects (950 per arm) will provide approximately 80% power to

demonstrate non-inferiority of PCI to CABG.

If non-inferiority is met, superiority testing will be performed with a one-sided alpha of 0.025.

The null ( $H_0$ ) and alternative ( $H_A$ ) hypotheses for the superiority test are:

$$H_0: F_{\text{PCI-PE}}(T) - F_{\text{CABG-PE}}(T) \geq 0$$

$$H_A: F_{\text{PCI-PE}}(T) - F_{\text{CABG-PE}}(T) < 0.$$

Using a one-sided alpha of 0.025, assuming 8% lost to follow-up at 3 years, the trial will have approximately 80% power to demonstrate superiority with a difference of 3.84% of PCI to CABG (e.g. 7.16% in the PCI arm vs. 11% in the CABG arm).

## 2.5.2 Major Powered Secondary Endpoints:

The first major powered secondary endpoint is defined as the composite of all cause death, MI, or stroke ( $\text{mRS} \geq 1$  and increase by  $\geq 1$  from baseline) occurring up to 30 days post randomization. It will be evaluated using the difference in Kaplan-Meier rates tested using the Com-Nougue approach<sup>1</sup> which is similar to the normal approximation to the binomial. The null ( $H_0$ ) and alternative ( $H_A$ ) hypotheses for non-inferiority of this major secondary endpoint is:

$$H_0: F_{\text{PCI-PE30}}(T) - F_{\text{CABG-PE30}}(T) \geq \Delta_{\text{PE30}}$$

$$H_A: F_{\text{PCI-PE30}}(T) - F_{\text{CABG-PE30}}(T) < \Delta_{\text{PE30}}.$$

$F_{\text{PCI-PE30}}(T) - F_{\text{CABG-PE30}}(T)$  is the difference in Kaplan-Meier rates at 30 days between PCI and CABG arms and  $\Delta_{\text{PE30}}$  is the non-inferiority margin for this powered secondary endpoint.

The power calculation is based on the following assumptions:

- composite event rate is 3% in each treatment arm at 30 days
- non-inferiority margin  $\Delta_{\text{PE30}} = 2\%$
- one-sided alpha = 0.05

A sample size of 1,900 subjects (950 per arm) will provide approximately 80% power to demonstrate non-inferiority of PCI to CABG.

The second major powered secondary endpoint is defined as the composite of all cause death, MI, stroke, or unplanned revascularization for ischemia occurring by 3 years post randomization. This composite will be evaluated using the difference in Kaplan-Meier failure rates at 3 years between CABG and PCI in all 3 patient populations. The hypothesis is designed to show non-inferiority of PCI to CABG with a one-sided alpha of 0.05. The null ( $H_0$ ) and alternative ( $H_A$ ) hypotheses for non-inferiority of this powered secondary endpoint are:

$$H_0: F_{\text{COMP-PCI}}(T) - F_{\text{COMP-CABG}}(T) \geq \Delta_{\text{COMP}}$$

$$H_A: F_{\text{COMP-PCI}}(T) - F_{\text{COMP-CABG}}(T) < \Delta_{\text{COMP}}.$$

$F_{\text{COMP-PCI}}(T) - F_{\text{COMP-CABG}}(T)$  is the difference in failure rates at 3 years between CABG and PCI and  $\Delta_{\text{COMP}}$  is the non-inferiority margin for this powered secondary endpoint.

The power calculation is based on the following assumptions:

- composite event rate is 22% in each treatment arm at 3 years
- minimum time to follow-up is 2 years
- median time to follow-up is approximately 3 years
- 8% lost to follow-up at 3 years
- non-inferiority margin  $\Delta_{\text{COMP}} = 8.4\%$
- one-sided alpha = 0.05
- accrual time of 29 months

A sample size of 1,900 subjects (950 per arm) will provide approximately 99% power to demonstrate non-inferiority of PCI to CABG using the Z statistic from the Com-Nougue approach<sup>1</sup>, which is similar to the normal approximation to the Binomial.

## 2.6 JUSTIFICATION OF DELTA

The non-inferiority margin of 4.2% for the primary endpoint is for the difference between the cumulative event rates at 3 years. For the PCI arm to pass the non-inferiority test for the primary endpoint of the composite of death, MI and stroke at 3 years, the maximum allowable event rate would be approximately 12.1%. Compared to the 11% event rate of the CABG arm, the average difference per year is approximately 0.4%. The criteria for an acceptable non-inferiority delta was carefully considered by the principal investigators, executive committee, PCI and surgical committees and country leaders of this protocol, representing more than 100 physicians not related to the study sponsor, 50% of whom are interventional cardiologists and 50% of whom are cardiac surgeons. A non-inferiority margin of 4.2% for the primary endpoint in this protocol has been agreed upon by this balanced study leadership to represent clinical therapeutic interchangeability between PCI and CABG, given the substantially lower peri-procedural morbidity of PCI, the likelihood for fewer strokes with PCI, especially in the first 30 days to 1 year (which in most cases is a clinically more important endpoint than MI, although the trial will not be powered to demonstrate a reduction in stroke), and the likely higher rate of subsequent unplanned revascularization for PCI.

## 2.7 STUDY SUCCESS

Study success is defined as passing the non-inferiority test of PCI to CABG on the primary endpoint of all cause death, MI or stroke ( $mRS \geq 1$  and increase by  $\geq 1$  from baseline) at 3 years. Detail of the test is specified in sections 2.5.1.

## **2.8 RANDOMIZATION AND BLINDING**

Approximately 1,900 subjects will be enrolled in a 1:1 ratio to either PCI with the XIENCE stent or CABG treatment. Randomization will be stratified by the presence vs. absence of medically treated diabetes, SYNTAX<sup>(4)</sup> score <23 vs.  $\geq 23$ , and study center. A centralized randomization service will be used.

Randomization will be performed after informed consent has been obtained and all eligibility criteria have been confirmed. Once randomized, the subject is considered registered in the trial and analyzed as part of ITT population. Once randomization is completed and a treatment is assigned, crossover is not permitted.

This is an unblinded clinical study.

### 3 ANALYSIS CONSIDERATIONS

#### 3.1 GENERAL STATISTICAL METHODS

Baseline demographic, clinical, angiographic, procedural and device data, and treatment results will be summarized using descriptive summary statistics. All data collected will be summarized overall and by treatment arms.

#### 3.2 Analysis of the Primary Endpoint

The primary analysis for non-inferiority and superiority will use the Kaplan-Meier estimates defined from randomization date. A sensitivity analyses for superiority will also be performed where time is defined from randomization. **Additional sensitivity analysis for the same non-inferiority and superiority test mentioned above will be rerun changing the start date from randomization to procedure date.** The stop date for all analyses defined at 3 years will be 1095 days from randomization.

Primary and supplemental analyses are defined below.

1. Primary Analysis of the Primary Endpoint for non-inferiority: Code for the primary analysis testing for non-inferiority with 1-sided alpha = 0.025, the non-inferiority margin  $\delta = 4.2\%$ .
  - Run the product-limit model:  
Proc lifetest data = DATAXX method = pl cs=none atrisk  
timelist = 1095 outsurv = OUTSURV ;  
Time Timevar\*Censvar(0);  
strata/ Group = treatment test = logrank;  
Run;
  - Use results in OUTSURV to compute formula in 2.3.1. Reject  $H_0$  if the critical value  $< -Z_{1-\alpha}$
2. Primary Analysis of the Primary endpoint using Z test for Superiority: Code for the primary endpoint testing for superiority with 1-sided alpha = 0.025: same as above, only use  $\delta = 0$  in calculations. This will be the primary test for superiority.
3. Sensitivity Analysis using Cox Model for Superiority: The primary endpoint is tested for superiority with 1-sided alpha = 0.05 using a Cox proportional hazards model. The proportional hazard assumption will first be tested by plotting the log-negative-log survival curves vs the log of survival time for each level of treatment group.  
  
If it is determined that the proportional hazard assumption is violated, then a logistic regression will be performed adjusting for time of follow-up in each group.  
  
Code to Test Proportional Hazard Assumption:  
proc lifetest data=DATAXX plot=(lls) noprint;  
time Timevar\*Eventvar(0);  
strata Group;  
run;

Code for Cox model where Proportion Hazard Assumption is **Not** violated:

```
ods output ParameterEstimates=Out1;  
Proc phreg data = DATAXX ;  
    model Timevar*Eventvar (0) = Group / rl;  
run;
```

Summary tables for time to event endpoints will include failure rates (Kaplan-Meier estimates), unadjusted hazard ratios, confidence interval for the hazard ratio, and a p-value.

Code for Logistic Regression used when Proportional Hazard Assumption is violated:

```
proc logistic data= DATAXX descending;  
    class Group;  
    model Event (event=1) =TimeVar Group / rl;  
run;
```

Where Event = all first events from randomization or procedure date to 1095 days  
Timevar = Days from Randomization or procedure date to last data point available for patient while on study. If the patient did not drop out prior to the analysis time point then this will be set to the analysis time point.

The odds ratio, 95% confidence interval of the odds ratio and p-value will be presented from this logistic model

### 3.3 Analysis of the Powered Secondary Endpoints

- The first powered secondary endpoints, the KM estimate of the composite rate of all-cause mortality, MI or stroke ( $mRS \geq 1$  and increase by  $\geq 1$  from baseline) from randomization to 30 days post randomization, will be analyzed in a similar fashion as the primary analysis of the primary endpoint described in Section 3.2 above. The key differences are that a 1-sided alpha is set at 0.05 and only non-inferiority will be tested at a margin of 2%.
- The second powered secondary endpoint is the KM estimate of the composite rate of all cause mortality, MI, stroke ( $mRS \geq 1$  and increase by  $\geq 1$  from baseline) or unplanned revascularization at 3 years. The same analysis will be performed as described for the primary analysis of the primary endpoint except the non-inferiority margin = 8.4%

### 3.4 Analysis of the Secondary Endpoints (Not powered)

All non-powered secondary endpoints will use a 2-sided alpha of 0.05.

#### 3.4.1 Continuous Variables

For continuous variables (e.g., age, percent diameter stenosis, and lesion length), results within treatment arm will be summarized with the numbers of observations, means, medians, standard deviations, 25th and 75th percentiles, minimums, and maximums per the table mockups.

Differences between the treatment arms, where specified, will be summarized with the differences of the two means, 95% confidence intervals for the difference between the means, and p-values based on a t-test. The distributions within each group will be tested for normality using the Shapiro-Wilks test and if normality cannot be assumed then a Wilcoxon rank-sum test and 95% confidence interval of the median will be presented. The confidence interval for the difference of two means will be calculated under the assumption of unequal variances.

Formulas for calculation of the confidence intervals for the continuous variables are given below:

1. 100(1- $\alpha$ )% Confidence Interval For The Difference of Two Means Under The Assumption of Unequal Variances Between The Two Groups<sup>5</sup>

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2}} SED$$

With the degrees of freedom for the approximate t statistic is determined by Satterthwaite's formula<sup>3</sup> as follows:

$$df = \frac{(w_1 + w_2)^2}{\frac{w_1^2}{n_1 - 1} + \frac{w_2^2}{n_2 - 1}}$$

where:

$\bar{x}_1$  = sample mean for group 1

$\bar{x}_2$  = sample mean for group 2

$s_1$  = sample standard deviation for group 1

$s_2$  = sample standard deviation for group 2

$n_1$  = sample size for group 1

$n_2$  = sample size for group 2

$$SED = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$



$$w_1 = \frac{s_1^2}{n_1}$$
$$w_2 = \frac{s_2^2}{n_2}$$

### 3.4.2 Categorical Variables

For categorical variables such as gender, in-hospital event rates, and angina status, results within treatment arm will be summarized with subject counts and percentages. Differences between the two treatment arms, where specified, will be summarized with the difference in percentages, the asymptotic 95% confidence interval for the difference of two percentages, and a p-value based on a chi-squared test. If 20% or more of the expected cell frequencies are less than 5 then a Fisher's Exact test will be used to test for differences in proportions.

For the determination of event rates in-hospital, the number of all subjects in the patient population will be used as the denominator. For variables ascertained at follow-up such as angina status, the denominator will be only those subjects who had follow-up performed at that time point. Unless otherwise noted, subjects with missing data are excluded from the denominator.

For secondary analyses of any type of death event for which event rates will be calculated, the denominator will include only subjects who have either had the event, died prior to the event time of interest, or had sufficient follow-up to be declared eligible for the endpoint. Sufficient follow-up implies that the subject had a follow-up visit past the lower limit of the window for the analysis time point (-7 days for the 30 day endpoint, and -30 days for yearly endpoints).

Formulas for calculating confidence intervals for the categorical variables are given below.

Please note that in using these formulas, it is assumed that the data are independent and binomially distributed. For the confidence interval for the difference of two proportions, it is also assumed that the two samples are independent, the samples sizes for the two groups are large ( $\geq 20$  per group), the expected event and non-event counts for both groups are at least 5, and the distribution is asymptotically normal.

1. 100(1- $\alpha$ )% Confidence Interval For The Difference Of Two Proportions<sup>6</sup>

$$\text{Lower Confidence Limit} = (\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}} \sqrt{\left( \frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2} \right) + \frac{1}{2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}$$

$$\text{Upper Confidence Limit} = (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}} \sqrt{\left( \frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2} \right) + \frac{1}{2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}$$

where:

$\hat{p}_1$  = sample proportion for group 1

$\hat{p}_2$  = sample proportion for group 2

$\hat{q}_1 = 1 - \hat{p}_1$

$\hat{q}_2 = 1 - \hat{p}_2$

$n_1$  = sample size for group 1

$n_2$  = sample size for group 2

$Z_{\frac{\alpha}{2}}$  = the (1 -  $\alpha/2$ ) Z - statistic

### 3.4.3 Time to Event Variables

Survival analysis techniques will be used to analyze the time-to-event variables that occur at or after 30 days of follow-up. All of these analyses will be performed with time defined from date of randomization and from date of procedure. Subjects without events will be censored at their last known event-free time point. If this event-free time point occurs after the analysis time point, the days to event variable will be set equal to the analysis time point so that the patient will be included in the analysis (e.g. if the last data point was collected at 3 years and 1 week post randomization, for the 3 year analysis, this patient is censored at exactly 3 years [1095 days]). For patients who did not have an event or early withdrawal and have not yet completed the 3 year visit, they will be censored at the time of their last follow-up. Time to first event curves will be constructed using Kaplan-Meier estimates.

For time to event variables such as MI, stroke and and revascularization, at all post discharge results will be summarized with Kaplan-Meier estimates of event rates. For time to event variables such as death and composite endpoints including death, deaths will be included for patients who dropped out early, but had a known death date, prior to study day 1095, from the national death registry.

For time to event analyses, hazard ratios, confidence interval for the hazard ratios, and p-values may also be presented from a Cox proportional hazards model.

Formulas are given below. Please note that in using these formulas, it is assumed that the data are independent observations.

The Kaplan-Meier product-limit estimator is defined as follows:

$$\hat{S}(T) = \begin{cases} 1 & \text{if } T_{\min} > T \\ \prod_{A \leq T_i \leq T} \left[ 1 - \frac{d_i}{r_i} \right] & \text{if } T_{\min} \leq T. \end{cases}$$

The variance of  $\hat{S}(T)$  is estimated by Greenwood's formula where

$$\hat{V}[\hat{S}(T)] = \hat{S}(T)^2 \sum_{A \leq T_i \leq T} \frac{d_i}{r_i(r_i - d_i)}$$

The Kaplan-Meier failure rate estimator and its variance are defined as

$$\begin{aligned} \hat{F}(T) &= 1 - \hat{S}(T) \\ \hat{V}[\hat{F}(T)] &= \hat{V}[\hat{S}(T)] \end{aligned}$$

100(1- $\alpha$ )% Confidence Interval for the hazard ratio

Cox proportional hazards regression will be used to obtain the confidence interval for the hazard ratio. The phreg procedure in SAS will be used to obtain the Wald confidence limits. An example of the code is as follows:

```
ods output ParameterEstimates=Out1;  
proc phreg data=InFile;  
  model TimeVar*Eventvar (0)= Group / rl;  
run;
```

### 3.4.4 Hypothesis Testing

Formal non-inferiority and superiority tests are planned for the primary and powered secondary endpoints. These endpoints are Kaplan-Meier failure rate estimates (higher rate is worse).

All non-inferiority tests will be one-sided. The null and alternative hypotheses will be of the following form (assuming showing A non-inferior to B):

$$\begin{aligned} H_0: \text{Endpoint}_A - \text{Endpoint}_B &\geq \Delta \\ H_A: \text{Endpoint}_A - \text{Endpoint}_B &< \Delta \end{aligned}$$

For one-sided superiority test, the null and alternative hypotheses will be of the form (assuming showing A superior to B):

$$\begin{aligned} H_0: \text{Endpoint}_A - \text{Endpoint}_B &\geq 0 \\ H_A: \text{Endpoint}_A - \text{Endpoint}_B &< 0 \end{aligned}$$

For two-sided superiority test, the null and alternative hypotheses will be of the form

$$H_0: \text{Endpoint}_A - \text{Endpoint}_B = 0$$

$$H_A: \text{Endpoint}_A - \text{Endpoint}_B \neq 0$$

Methods for testing are as follows. Detailed specifications on non-inferiority deltas, significance levels, and decision rules are provided in the section on analysis of endpoints.

1. Test for non-inferiority

The test for non-inferiority will be based on the one-sided  $(1-\alpha)$  confidence interval of the difference between the two treatment arms. If the upper bound of the one-sided  $(1-\alpha)$  confidence interval is less than  $\Delta$ , then it will be concluded that the null hypothesis can be rejected and non-inferiority will be declared.

2. Test for superiority

For one-sided superiority test, if the one-sided  $(1-\alpha)$  confidence interval is less than 0, then it will be concluded that the null hypothesis can be rejected and superiority will be declared.

For two-sided superiority test, if the  $(1-\alpha)$  confidence interval does not contain 0, then it will be concluded that the null hypothesis can be rejected and superiority will be declared.

Caution must be exercised when interpreting p-values displayed for analyses other than those performed for the primary and powered secondary endpoints, as the study was not powered to detect differences for any of those other variables. The resulting p-values, whether or not less than 0.05, may be a result simply due to pure chance.

### 3.5 SUBGROUPS FOR ANALYSIS

The primary and major powered secondary endpoints will be analyzed by the following subgroups. The treatment comparisons in these analyses are not powered for hypothesis testing and are descriptive in nature. Data described below will be summarized as seen in the table mockups.

For each covariate, two models will be run:

- one with just treatment and the covariate. It will be testing if there is a significant effect regardless of the the covariate. The 3 year Kaplan-Meier estimates for each treatment along with the percent difference and 95% confidence interval of the differences will be presented, as well as hazard ratios and 95% confidence intervals of the hazard ratios. Forest plots will be presented, including results for all subgroups analyzed
- a second model, containing also the interaction term of covariate and treatment, will be testing if the treatment effect varies significantly with the covariate values. For any statistically significant interaction terms a Kaplan-Meier plot will be displayed by subgroup.

Subgroup details are as follows.

**Diabetic subgroups:**

- Diabetes mellitus requiring medication, defined as subjects treated with oral hypoglycemic medications or insulin vs. non treated or no diabetes

**Age:**

- Age  $\geq$  Median, Age  $<$  Median
- Age  $\geq 75$ , Age  $< 75$

**Gender:** male vs female

**Other baseline characteristics:**

- **Body mass index:**  $\geq$  Median,  $<$  Median
- **Prior MI:** yes vs. no
- **LVEF**  $\geq$  Median, LVEF  $<$  Median
- **LVEF**  $> 40\%$ , LVEF  $\leq 40\%$
- **Chronic kidney disease (CKD):** Creatinine Clearance (ml/min) from Cockcroft-Gault formula:  $[(140 - \text{age})(\text{weight in kg}) / (72 \times \text{Serum creatinine in ml/min})] (*0.85 \text{ in women})$   
- CKD  $<$  or  $\geq 60 \text{ ml/min}$
- **Geographic location** (US vs. EU vs. Other); US vs other; NA vs. EU vs other; NA vs. other

**Disease characteristics:**

- **Number of diseased vessels**  $\geq 3$ ,  $< 3$ ;  $\geq 2$ ,  $< 2$  (core lab)
- **Distal Left Main Bifurcation Involvement** (yes/no) (core lab)
- **Chronic Total Occlusion of non left main** (yes/no) from angiography (core lab)

Scores at baseline:

- **SYNTAX score:**
  - by core lab  $\geq$  Median,  $<$  Median; and by tertiles
  - by core lab  $\geq 23$ ,  $< 23$
  - by core lab  $\geq 33$ ,  $< 33$
  - clinical SYNTAX score  $\geq$  Median,  $<$  Median
  - clinical SYNTAX score  $\geq 23$ ,  $< 23$
  - clinical SYNTAX score  $\geq 33$ ,  $< 33$

- Syntax Score II (SSII); developed by Farooq et al.<sup>7</sup>  $\geq$  Median,  $<$  Median; and by tertiles
- ACEF score = age (years)/ejection fraction (%) (+1 if serum creatinine value is  $>2$  mg/dL), then categorized as  $<$  or  $\geq$  median

### 3.6 EXPLORATORY ANALYSIS

#### 3.6.1 Exploratory Composite Endpoint Analysis

##### 3.6.1.1 Weighted Composite Endpoint

Exploratory analysis comparing PCI and CABG on the weighted composite endpoint of all cause death, MI or stroke ( $mRS \geq 1$  and increase by  $\geq 1$  from baseline) occurring up to 3 years post randomization will be performed. The weighting algorithm will be based on that developed for death, MI, and stroke by Tong et al.<sup>8</sup>.

The 3 year Kaplan Meier event rates will be calculated for columns 2 and 3 below as well as the difference between the two. The relative weights are from the Tong reference. The normalized relative weight is calculated as follows:

Normalized Relative Weight = (Relative Weight for event/Relative Weight Sum) x Number of different types of events

Adjusted Difference = Difference x Normalized Relative Weight

Event	KM estimate in CABG by 3 years	KM estimate in PCI by 3 years	Difference (PCI-CABG)	Relative Weight from Tong et al	Normalized Relative Weight from Tong et al	Adjusted Difference
Death				0.23	1.25	
Stroke				0.18	0.98	
MI				0.14	0.76	
Composite				0.55		

The adjusted difference will then be placed into the formula for the Z statistic in Section 2.3.1 and p-values and 95% confidence intervals will be displayed from this.

##### 3.6.1.2 The Win Ratio

Exploratory analysis of comparing PCI and CABG using the win ratio calculated by Pocock et al method<sup>8</sup> with the unmatched approach<sup>7</sup> will be performed. The outcomes used in calculating the win ratio ordered by their clinical importance (from high to low) will be all-cause death (high), stroke ( $mRS \geq 1$  and increase by  $\geq 1$  from baseline) and MI (low). Here are the steps needed to follow for this calculation:

Let  $N_n$  = total patients on PCI

Let  $N_s$  = total patients on CABG

Compare the timing of death, stroke and MI for each patient on PCI to each patient on CABG ( $N_n \times N_s$  comparisons) for the shortest time duration that exists between each pair of patients.

Classify each patient into one of 7 categories:

- (a) PCI patient died prior to CABG patient
- (b) CABG patient died prior to PCI patient
- (c) PCI patient had stroke prior to CABG patient
- (d) CABG patient had stroke prior to PCI patient
- (e) PCI patient had MI prior to CABG patient
- (f) CABG patient had MI prior to PCI patient
- (g) None of the above

Since groups b, d and f are in favor of PCI,  $N_b + N_d + N_f = N_{win}$

Since groups a, c and e are not in favor of PCI,  $N_a + N_c + N_e = N_{loss}$

Win Ratio ( $R_w$ ) =  $N_{win}/N_{loss}$

A p-value will be calculated for the Finkelstein-Schoenfeld test<sup>10</sup>, as well as a 95 % CI interval will be calculated using bootstrapping methods<sup>11</sup>.

### 3.6.1.3 Covariate Adjusted Analysis

A multivariable Cox model (or logistic regression if the proportional hazards assumption is not met) will be performed, adjusting for baseline variables historically known to be prognostically important in order to identify independent correlates for the outcomes of interest, as follows (number of variables selected to avoid model over-fitting; i.e. ~1 variable for every 10 events based on the anticipated rates):

**Outomce = All-cause death/MI/stroke up to 30 day:**

- age, gender, diabetes, LVEF, SYNTAX score groups (core lab), randomization arm (forced in)

**Outcome= All-cause death/MI/stroke up to 3-year:**

- age, gender, diabetes, hypertension, hyperlipidemia, prior MI, current smoking, prior cerebrovascular disease, prior PAD, prior heart failure, prior COPD, presentation in ACS vs stable CAD, baseline anemia, baseline creatinine clearance, LVEF, SYNTAX score groups (core lab), randomization arm (forced in)

**Outcome= All-cause death/MI/stroke/unplanned revascularization up to 3-year:**

- same as all all-cause death/MI/stroke

Continuous data such as age will be entered as continuous except as otherwise noted.

For Syntax score groups: <23, 23-32, >32, two dummy variables will need to be created:

If Syntax score < 23 then Syntax23=1;  
Else if Syntax ne . then Syntax23 = 0;  
If Syntax score > 32 then Syntax32=1;  
Else if Syntax ne . then Syntax32 = 0;  
Add both of the above dummy variables to the model in place of the 3 way Syntax I score.

Appropriate methods may be used to impute missing baseline values.

1). Test the proportional hazard assumption:

The proportional hazard assumption will first be tested by plotting the log-negative-log survival curves vs the log of survival time for each level of treatment group against each covariate one at a time.

Code to Test Proportional Hazard Assumption:

```
proc lifetest data=DATAXX plot=(lls) noprint;  
time Timevar*Eventvar(0);  
strata Covariate;  
run;
```

2). If it is determined that the proportional hazard assumption is NOT violated then a Cox regression model will be performed with a stepwise option.

Code for Cox model where Proportion Hazard Assumption is **Not** violated:

```
Proc phreg data=infile;  
Model timevar*Eventvar(0)=group baselinevars ;  
Run;
```

Summary tables from this model will include failure rates (Kaplan-Meier estimates), adjusted hazard ratios, confidence interval for the hazard ratio, and a p-value.

3). If it is determined that the proportional hazard assumption is violated then a Logistic regression model will be performed adjusting for time on study with a stepwise option.

Code for Logistic Regression used when Proportional Hazard Assumption is violated:

```
proc logistic data= DATAXX descending;  
class Group;  
model Event (event=1) =TimeVar Group baselinevars/ rl;  
run;
```

Where Event = all first events from randomization or procedure date to 1095 or 30 days (dependent on outcome)

Timevar = Days from Randomization or procedure date to last data point available for patient while on study. If the patient did not drop out prior to the analysis time point then this will be set to the analysis time point.



The odds ratio, 95% confidence interval of the odds ratio and p-value will be presented from this logistic model

#### **3.6.1.4 Other Exploratory Analyses**

Utility of the SYNTAX score, SYNTAX score II, ACEF score, clinical SYNTAX score, and novel predictive instruments will be investigated.

### **3.7 FOLLOW-UP TIME POINTS**

All subjects will have follow-up as a telephone contact or office visit at  $30 \pm 7$  days,  $180 \pm 14$  days, and at 1, 2, 3, 4 and 5 years (-30days/+60 days). Procedures and tests will be performed at each follow-up time point as per the Schedule of Events within the study protocol. Subjects who have not yet completed the 3-year visit (unless the most recent follow-up was completed within 28 days of the AE check initiation), or who completed the 3 year visit before the lower bound of the 28 day window (i.e. before 1067 days after randomization) will be followed up with a phone call or office visit prior to database snapshot to capture potential withdrawal dates and any events that may have occurred since the last visit. Events dates and early withdrawal date will be collected at this call.

Primary analyses for the study will take place when approximately 50% of the patients have reached the 3-year follow-up AND all patients have reached the 2-year follow-up, whichever time point is latest.

### **3.8 HANDLING OF MISSING DATA**

Every effort will be made to minimize the loss of subjects during this study. Subjects discontinuing early from the study and their reason for dropping out will be summarized by group to see if any biases in the analyses may have been created by any differences between groups.

The impact of missing data will be minimized in the analysis of the primary and key secondary endpoints by performing the analyses using Kaplan-Meier estimates, where subjects will be censored at their last assessed time point.

### **3.9 ASSESSMENT OF DATA POOLABILITY**

#### **3.9.1 Multiple Center Effect**

Analyses will be performed pooling data across study sites.

The EXCEL trial will have approximately 165 sites in USA, Canada, Europe, South America, South Korea and Australia, where two treatment strategies (PCI vs CABG) will be compared. Descriptive analysis (number and rate of primary endpoint events) will be performed to assess the consistency of results across sites. For the analysis of center effect, data from smaller sites may be combined by region for the analysis.

#### **3.9.2 Multiple Region Effect**

Analyses will be performed pooling data across study regions.

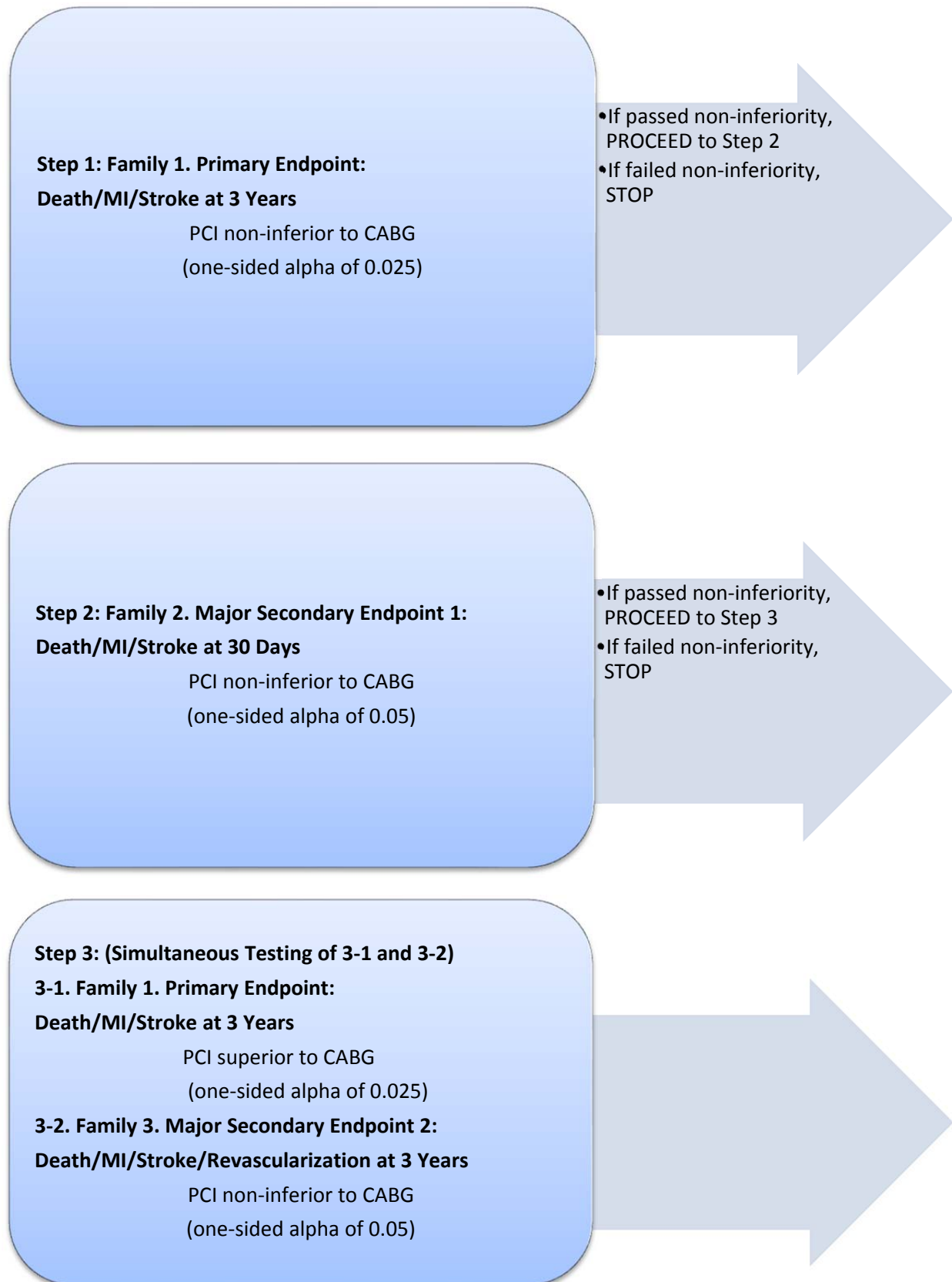
Data will be presented by region, with region being defined by the US, Europe, and outside of US/Europe. Data will also be presented as North America (US/Canada), Europe and outside of NA/EU. To evaluate the potential differences in treatment effect for different regions for the primary and powered secondary endpoints, the interaction p-value from the Cox regression model including region, treatment, and region by treatment interaction term will be compared against a significance level of 0.15.

### **3.9.3 Multiple Study Stent Effect**

The two study stents are considered to be interchangeable in the study and the EXCEL trial is designed to evaluate two treatment strategies (PCI and CABG) instead of any specific study device. To provide evidence for poolability between the two study stents, bioequivalence analysis showing that XIENCE PRIME and XIENCE V are bioequivalent have been submitted to the FDA in the XIENCE PRIME PMA submission (P110019). XIENCE PRIME, XIENCE Xpedition, and XIENCE PRO are the same stent. Descriptive analysis will be performed to assess the consistency of results between the study stents.

### **3.10 MULTIPLICITY ISSUES**

If EXCEL has met the primary analyses of its primary endpoint of the non-inferiority test on the ITT population, additional hypothesis tests for the of the powered secondary endpoints based on the ITT population will be performed following a fixed sequence as below (FWER=family wise error rate). All of these analyses will be performed from time from randomization:



For hypothesis test(s) of each of the major powered secondary endpoints, the type I error is a 1-sided alpha of 0.05. If any hypothesis testing fails in steps 1 and 2, no further testing for the following step will be performed for labeling purposes. However they will still be tested for research and publication purposes. For example, if PCI failed to show non-inferiority to CABG for the major power secondary endpoint of Death/MI/Stroke at 30 Days (step 2), then the hypothesis in step 3 will not be tested for labeling purposes, but may be tested for exploratory reasons. Caution must be exercised when interpreting p-values displayed for analyses other than those performed for the primary and key powered secondary endpoints, as the study was not powered to detect differences on any of those other variables. The resulting p-values, whether or not less-than 0.05, may be a result simply due to chance, and are displayed for hypothesis-generating purposes only.

### **3.11 INTERIM ANALYSIS**

No formal interim analyses are planned for this study. As such, no formal statistical rule for early termination of the trial is defined. Interim study reports with descriptive analysis may be produced for regulatory or reimbursement purposes.

### **3.12 DOCUMENTATION AND OTHER CONSIDERATIONS**

All analyses will be performed using SAS<sup>®</sup> for Windows, version 9.1<sup>5, 12</sup> or higher.

## **4 VARIABLES FOR ANALYSIS**

This section describes the additional variables to be presented. In general, all analyses will be summarized overall and by treatment group, unless otherwise noted in the text below or in the table mockups.

### **4.1 ENROLLMENT AND SUBJECT DISPOSITION**

The count of subjects enrolled at each study site will be summarized. Subjects terminating from the study by 30, 180 days, and 1, 2, 3, 4 and 5 years, and the reason for early termination, will be summarized with subject counts and percentages.

### **4.2 DEMOGRAPHICS, SUBJECT CHARACTERISTICS, AND PROCEDURE INFORMATION**

Demographic, baseline subject characteristic, and procedure information data will be summarized overall, by treatment group, and for the difference between treatment groups using descriptive statistics, as per the table mockups. The data will be summarized for each analysis population to show the balance between the treatment arms in support of the various endpoint analyses. The variables for analysis are given in each section below.

#### **4.2.1 Demographics**

- Age (in years)
- Gender at Birth

#### **4.2.2 Risk Factors**

The presence of the following risk factors:

- Current Tobacco Use
- All Diabetes Mellitus
- Non-Diabetes Mellitus
- Diabetes Mellitus Requiring Medication
- Diabetes Mellitus Requiring Insulin
- Non-medically Treated Diabetics
- Hypertension Requiring Medication
- Hypercholesterolemia Requiring Medication

#### **4.2.3 Cardiac History**

The presence of the following cardiac histories and classification of the disease:

- Prior MI
- MI within 2 Months
- History of Angina (stable angina, unstable angina)
- All Prior Cardiac Interventions
  - Prior PCI

- Prior CABG
- Prior Cardiac Intervention of Left Main

#### **4.2.4 Physical Measurements**

- Height (cm)
- Weight (kg)
- Body Mass Index (BMI) (kg/m<sup>2</sup>)
- Systolic Blood Pressure (mmHg)
- Diastolic Blood Pressure (mmHg)

#### **4.2.5 Disease State/Severity**

- SYNTAX Score
- SYNTAX Score II
- Clinical SYNTAX Score
- ACEF Score
- LVEF
- Chronic obstructive lung disease
- Peripheral vascular disease
- Prior stroke
- Modified Rankin Scale (mRS) for stroke
- Number of Diseased Vessels
- Number with  $\geq 2$  Diseased Vessels
- Number with  $\geq 3$  Diseased Vessels
- Number of Vessel Treated
- Number with  $\geq 3$  vessels Treated
- Number with  $\geq 2$  vessels Treated
- Distal left main bifurcation involvement
- Chronic total occlusion present

#### **4.2.6 Laboratory Tests**

- HgbA1c
- hsCRP
- BNP
- Hgb
- WBC
- Platelets
- CK-MB
- Serum Creatinine

#### **4.2.7 Procedure/Device Information**

Procedure and device information will be summarized for the two treatment arms separately. No comparisons will be made.

#### **4.2.7.1 Procedure/Device Information for PCI arm only**

The following device information will be presented in a listing by-subject:

- GPIIb/IIIa Inhibitor Usage During Procedure
- Antiplatelet Usage During Procedure
- Maximum Balloon Pressure Used Over the Entire Procedure (pre-dilatation, stent deployment, post-dilatation)
- Total Number of Stents Placed
- Total Length of All Stents Placed (mm)
- IVUS and/or FFR guidance
- distal left main vs. left main ostial/body
- treatment strategies for the distal left main bifurcation
- chronic total occlusions
- bifurcation lesions

The following device information data will be summarized on a by-stent basis:

- Diameter of Stents Used (mm)
- Length of Stents Used (mm)

#### **4.2.7.2 Procedure Information for CABG arm only**

The following device information will be summarized in a listing by-subject:

- Total Number of grafts placed
- On pump vs. Off pump
- Single vs. bilateral ITA vs. multiple arterial graft use
- Endoscopic versus open saphenous vein harvest technique
- Epi-aortic ultrasound and/or TEE
- Prophylactic and management strategies for atrial fibrillation
- Carotid Screening

#### **4.2.7.3 Device Malfunction**

Device malfunctions and their outcomes will be listed for the PCI arm only, using subject counts.

- Device Malfunction
- Outcome of Malfunction

#### **4.2.8 Protocol Medications**

Antiplatelet use at discharge, 30 days, 180 days, and 1, 2, 3, 4 and 5 years summarized as:

- Number of Subjects on Clopidogrel/Ticlopidine/Prasugrel/Ticagrelor
- Duration (days) on Clopidogrel/Ticlopidine/Prasugrel/Ticagrelor
- Number of Subjects on Aspirin
- Duration (days) on Aspirin
- Number of Subjects on Dual Antiplatelet Therapy (both aspirin and Clopidogrel/Ticlopidine/Prasugrel/Ticagrelor)
- Duration (days) on Dual Antiplatelet Therapy

#### **4.2.9 Angina Status**

- Angina Status and classification at discharge, 30 days, 180 days, 6 months, and 1, 2, 3, 4 and 5 years

### **4.3 MORPHOLOGY AND QUANTITATIVE CORONARY ANGIOGRAPHY**

Lesion morphology and quantitative coronary angiography (QCA) data at the target lesion(s) will be summarized overall, by treatment group, and for the difference between treatment groups using descriptive statistics, as per the table mockups. The variables to be summarized are listed in the sections below.

#### **4.3.1 SYNTAX Score**

- SYNTAX score at baseline.

#### **4.3.2 Morphology**

Pre-procedure (all subjects):

- Target Lesion Vessel (CASS Site Location)
- Thrombus
- Aneurysm
- Calcification
- Eccentric Lesion
- Lesion Angulation > 45°
- American College of Cardiology/American Heart Association (ACC/AHA) Lesion Class

Post-procedure (PCI arm only):

- Thrombus
- Aneurysm
- Dissection

#### **4.3.3 Quantitative Coronary Angiography**

Note that for the variables listed below, the reference vessel diameter (RVD) will be calculated using the user-defined method and percent diameter stenosis will be calculated using the interpolated method.

Pre-procedure (all subjects):

- Reference Vessel Diameter (RVD, in mm)
- Lesion Length (mm)
- Minimal Lumen Diameter (MLD, in mm)
- Percent Diameter Stenosis (%DS)

Post-procedure (PCI arm only):



- Reference Vessel Diameter (RVD, in mm)
- Minimal Lumen Diameter (MLD, in mm, In-Stent, In-Segment, Proximal, and Distal
- Percent Diameter Stenosis (%DS), In-Stent, In-Segment, Proximal, and Distal

#### **4.4 SAFETY ANALYSIS**

Adverse event and treatment results will be summarized for the ITT population only. AEs will be summarized by severity and relatedness. Serious AEs will be presented in a listing.

## 5 Changes From Protocol

At the request of the FDA reviewer, the As Treated (AT) analysis set is added for the analysis of the primary and secondary powered endpoints.

For consistency purposes, for the 30-day time-point, the primary endpoint will be measured as a KM estimate (as it is being done for all other time-points), and not as a rate. As such, it will be analyzed with the same methods as this endpoint is analyzed at the other time-points.

## 6 REFERENCES

- (1) Com-Nougue C, Rodary C, Patte C, How to establish equivalence when data are censored: a randomized trial of treatments for B non-Hodgkin lymphoma. *Statist. Med.* 1993. 12, 1353–1364.
- (2) Morice MC, Serruys PW, Kappetein AP, et al. Outcomes in patients with Institute, Cary, NC, 2005. De novo left main disease treated with either percutaneous coronary intervention using paclitaxel-eluting stents or coronary artery bypass graft treatment in the Synergy Between Percutaneous Coronary Intervention With TAXUS and Cardiac Surgery (SYNTAX) Trial. *Circulation.* 2010; 121:2645-2653.
- (3) Hintze J. NCSS and PASS. Number Cruncher Statistical Systems. Kaysville, Utah. 2005.
- (4) Sianos, Georgios et al., The SYNTAX Score: an angiographic tool grading the complexity of coronary artery disease. *EuroInterv.* 2005;1:219-227.
- (5) SAS/STAT User's Guide, Version 9.1, SAS Publishing, 2004.
- (6) Stokes ME, Davis CS, Koch GG. Categorical Data Analysis Using The SAS System. SAS Institute, Inc. Cary, NC, 2000.
- (7) Farooq V, et al., Anatomical and clinical characteristics to guide decision making between coronary artery bypass surgery and percutaneous coronary intervention for individual patients: development and validation of SYNTAX score II. *Lancet.* 2013. 381, 639-50.
- (8) Pocock SJ, et al., The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *European Heart Journal*, 2012. 33, 176-182.
- (9) Tong BC, et al., Weighting Composite Endpoints in Clinical Trials: Essential Evidence for the Heart Team. *Ann Thorac Surg*, 2012. 94, 1908-1913.
- (10) Finkelstein D.M. and Schoenfeld D.A., Combining mortality and longitudinal measures in clinical trials. *Statistics in Medicine*, 1999; 18(11): 1341-1354.
- (11) Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine.* 2000;19(9):1141-64.
- (12) Dmitrienko, A. et al., Analysis of Clinical Trials Using SAS: A Practical Guide, SAS

## 7 APPENDICES

### APPENDIX A: Planned Analyses for IFU

In addition to the pre-specified hypotheses testing on the primary and major secondary endpoints, descriptive analysis regarding the components of the primary and major secondary endpoints will be provided for the ITT population. Moreover, the primary and major secondary endpoint, along with their components will be provided for the following subsets of subjects:

- Diabetics
- Female
- Elderly (age  $\geq 65$  years)
- Number of diseased vessels (with  $\geq 2$  diseased vessels)
- Bifurcation in LM

All subgroups are defined based on information related to the index procedure. If a subgroup is defined by lesion level characteristics, then in general, the subgroup is restricted to single lesion treated subjects and subjects with staged procedure(s) will be excluded.

The following Quality of Life (QOL) scores/scales collected at baseline, 1 month, 6 months, 12 months, and 3 years will be summarized using mean and standard deviation:

- Seattle Angina Questionnaire (SAQ)
- London School of Hygiene Dyspnea Questionnaire
- Medical Outcomes Study 12-item Short Form Health Status questionnaire (SF-12)
- Patient Health Questionnaire (PHQ-8)
- EuroQol (EQ-5D)