



PROTOCOL ABT-CIP-10235

XIENCE 28 Global Study

Statistical Analysis Plan

[REDACTED]
June 19, 2020

[REDACTED]
an Zheng
Jin Wang

Confidential and Proprietary

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

TABLE OF CONTENTS

1. SYNOPSIS OF STUDY DESIGN AND PROCEDURES3

1.1 Purpose of the Statistical Analysis Plan.....3

1.2 Study Objectives.....3

1.3 Study Design.....3

 1.3.1 Selection of Control4

 1.3.2 Primary Endpoint5

 1.3.3 Secondary Endpoint(s).....5

1.4 Analysis Populations5

 Primary Analysis Population5

1.5 Sample Size Calculations6

1.6 Study Success.....7

2. ANALYSIS CONSIDERATIONS8

2.1 Statistical Methods8

 2.1.1 Descriptive Statistics for Continuous Variables8

 2.1.2 Descriptive Statistics for Categorical Variables10

 2.1.3 Propensity Score12

 2.1.3-1 Propensity Score Variable List.....13

 2.1.4 Hypothesis Testing.....14

 2.1.5 Survival Analyses.....15

2.2 Endpoint Analyses16

 2.2.1 Primary Endpoint Analysis16

 2.2.3 Secondary Endpoint Analyses16

2.3 Subgroups for Analysis.....16

 2.3.1 Sex17

 2.3.2 Diabetes17

 2.3.3 Other Subgroups17

2.4 Analysis Window.....17

2.5 Handling of Missing Data17

 2.6.1 Multiple Center Effect17

2.7 Adjustments for Covariates18

2.8 Multiplicity Issues18

2.9 Sensitivity Analysis.....18

2.10 Documentation and Other Considerations19

3. ACRONYMS AND ABBREVIATIONS20

4. REFERENCES.....21

[REDACTED]22

[REDACTED]22

[REDACTED]23

1. SYNOPSIS OF STUDY DESIGN AND PROCEDURES

1.1 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 Protocol ABT-CIP-10235, the XIENCE 28 Global clinical study. This plan is based on the [REDACTED] November 13, 2018 study protocol.

1.2 Study Objectives

The primary objective of this trial is to further evaluate safety of 1-month (as short as 28 days) dual antiplatelet therapy (DAPT) in subjects at high risk of bleeding (HBR) undergoing percutaneous coronary intervention (PCI) with XIENCE.

1.3 Study Design

XIENCE 28 Global Study is a prospective, single arm, multi-center, open label trial to further evaluate the safety of 1-month (as short as 28 days) dual antiplatelet therapy (DAPT) in subjects at high risk of bleeding (HBR) undergoing percutaneous coronary intervention (PCI) with the approved XIENCE family of coronary drug-eluting stents.

The XIENCE family stent systems include commercially approved^a XIENCE Xpedition Everolimus Eluting Coronary Stent System (EECSS), XIENCE Alpine EECSS, XIENCE PRO^X EECSS^b, XIENCE Pro^A EECSS^c or XIENCE Sierra EECSS which are all manufactured by Abbott Vascular, Inc, Santa Clara, USA. The above listed XIENCE stents will hereinafter be called “XIENCE” in this trial.

A minimum of 800 to a maximum of 960 subjects from approximately 50 sites globally will be registered in this trial. Subject registration is capped at 120 per site. Trial population consists of non-complex HBR subjects with up to three native coronary artery lesions (a maximum of two lesions per epicardial vessel) with reference vessel diameter between 2.25 mm and 4.25 mm. Eligibility of P2Y12 receptor inhibitor discontinuation will be assessed at 1-month follow-up. Subjects who are free from myocardial infarction (modified ARC), repeat coronary revascularization, stroke, or stent thrombosis (ARC definite/probable) within 1 month (prior to 1-month visit but at least 28 days) after stenting AND have been compliant with 1-month DAPT

^a For each geography included in the trial, only approved and commercial available XIENCE stent(s) in that geography will be used.

^b XIENCE PRO^X is a rebrand of the XIENCE Xpedition Stent System.

^c XIENCE PRO^A is a rebrand of the XIENCE Alpine Stent System.

without interruption of either aspirin and/or P2Y12 receptor inhibitor for > 7 consecutive days are considered as “1-month clear”, and will discontinue P2Y12 receptor inhibitor as early as 28 days and continued with aspirin monotherapy through 12-month follow-up.

All registered subjects will be followed at 1, 3, 6 and 12 months post index procedure.

The data collected from the XIENCE 28 Global Study will be compared with the historical control of non-complex HBR subjects treated with standard DAPT duration of up to 12 months from the XIENCE V USA Study, which is a US post-approval study to evaluate the safety of XIENCE V EECSS in “all-comer” population under real-world setting.

1.3.1 Selection of Control

The XIENCE V USA historical HBR control is derived based on the following criteria:

Definition of non-complex HBR from XIENCE V USA

- HBR inclusion criteria (any one of the below HBR criteria):
 - Age \geq 75 years
 - History of major bleeding
 - History of stroke
 - Receiving or scheduled to receive chronic anticoagulation therapy
 - Renal insufficiency (creatinine > 2mg/dl)
 - Anemia (Hb < 11g/dl or transfusion)
 - Thrombocytopenia (platelet count < 100.000/mm³)
- Exclusion criteria for non-complex:
 - STEMI
 - LVEF < 30%
 - Patients with more than 3 lesions treated during index procedure
 - At least one lesion with RVD < 2.25 mm or > 4.25 mm (visual estimation)
 - At least one lesion located in left main
 - At least one lesion located in graft
 - At least one CTO lesion
 - At least one in-stent restenosis lesion
 - At least one target lesion with length > 32 mm by visual estimation.

The above selection criteria for the XIENCE V USA historical control aligns with the key inclusion/exclusion criteria of the XIENCE 28 Global Study.

For primary analysis, the XIENCE V USA non-complex HBR control subjects must be also 1-month clear, following the same logic as defined for the primary analysis population (refer to Section 1.4).

1.3.2 Primary Endpoint

The primary endpoint is a composite of Net Adverse Clinical Endpoint (NACE, a composite rate of all-cause death, all myocardial infarction (modified Academic Research Consortium [ARC]), stent thrombosis (ARC definite or probable), stroke or major bleeding (Bleeding defined by the Bleeding Academic Research Consortium [BARC] type 2-5) from 1 to 6 months.

1.3.3 Secondary Endpoint(s)

The following endpoints will be assessed from 1 to 6 months:

- Stent thrombosis (ARC definite/probable, ARC definite)
- All death, cardiac death, vascular death, non-cardiovascular death
- All myocardial infarction (MI) and MI attributed to target vessel (TV-MI, modified ARC)
- Composite of cardiac death or MI (modified ARC)
- Composite of all death or all MI (modified ARC)
- All stroke, ischemic stroke and hemorrhagic stroke
- Clinically-indicated target lesion revascularization (CI-TLR)
- Clinically-indicated target vessel revascularization (CI-TVR)
- Target lesion failure (TLF, composite of cardiac death, TV-MI and CI-TLR)
- Target vessel failure (TVF, composite of cardiac death, TV-MI and CI-TVR)
- Bleeding defined by the Bleeding Academic Research Consortium (BARC) type 2-5 and type 3-5

The primary endpoints of NACE and all the above secondary endpoints will also be assessed from 6 months to 12 months and from 1 month to 12 months, respectively.

1.4 Analysis Populations

Primary Analysis Population

The primary analysis population includes “1-month clear” population, defined as subjects who are free from myocardial infarction (modified ARC), repeat coronary revascularization, stroke, or stent thrombosis (ARC definite/probable) within 1 month (prior to 1-month visit but at least 28

days) after stenting AND have been compliant with 1-month dual antiplatelet therapy (DAPT) without interruption of either aspirin and/or P2Y12 receptor inhibitor for > 7 consecutive days.

If a subject was event-free and DAPT compliant (as defined above) during the first month, but does not stop P2Y12 inhibitor after the 1-month visit, due to physician's or subject's decision, this subject will be considered as "1-month not clear".

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

In this SAP the term '1-month DAPT arm' refers to subjects who are from the XIENCE 28 Global Study only, unless otherwise specified.

1.5 Sample Size Calculations

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- Sample size in the control arm (1-month clear): ~1,400

[REDACTED]

A sample size of approximately 800 subjects in the 1-month DAPT arm will be required for this test.

[REDACTED]

[REDACTED]

1.6 Study Success

The trial success is defined as passing the non-inferiority test on the primary endpoint of NACE for the “1-month clear” population. Detail of the test is specified in section 2.1.4.

2. ANALYSIS CONSIDERATIONS

2.1 Statistical Methods

Baseline demographic, clinical, angiographic, procedural, and device data, and clinical results will be summarized using descriptive summary statistics.

2.1.1 Descriptive Statistics for Continuous Variables

For continuous variables (e.g., age, percent diameter stenosis and lesion length), results will be summarized with the numbers of observations, means, and standard deviations [REDACTED], with quartiles, minimums, maximums, and two-sided 95% confidence intervals for the means [REDACTED]. Differences between two comparison groups of interest, where specified, will be summarized with the differences of the two means, and two-sided 95% confidence intervals for the difference between the means. These calculations will be done under the assumption that the data for the two arms are independent and approximately normal in distribution. The confidence interval for the difference of two means will be calculated under the assumption of unequal variances. If the asymptotic assumptions fail, then nonparametric summary statistics (medians, 25th and 75th percentiles) may be displayed as an alternative.

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

1. 100(1- α)% Confidence Interval For A Single Mean⁵

$$\bar{x} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

where:

\bar{x} = sample mean

s = sample standard deviation

n = sample size

$t_{\frac{\alpha}{2}}$ = the alpha/2 t - statistic for n – 1 degrees of freedom

2. 100(1-α)% Confidence Interval For The Difference of Two Means Under The Assumption Of Equal Variances Between The Two Groups⁵

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2}} \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

where:

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

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

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

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

$t_{\frac{\alpha}{2}}$ = the alpha/2 t - statistic for $n_1 + n_2 - 2$ degrees of freedom

3. 100(1- α) % Confidence Interval for the Difference of Two Means under the Assumption of Unequal Variances between the Two Groups⁵

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

With the degrees of freedom for the approximate t statistic is determined by Satterthwaite's formula² 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

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

$$w_1 = \frac{s_1^2}{n_1}$$

$$w_2 = \frac{s_2^2}{n_2}$$

2.1.2 Descriptive Statistics for Categorical Variables

For categorical variables such as gender, NACE and TLF, results will be summarized with subject counts and percentages/rates, [REDACTED] with exact two-sided 95% Clopper-Pearson⁶ confidence intervals. Differences between two comparison groups of interest, when specified, will be summarized with the difference in percentages and the Newcombe⁵ score two-sided 95% confidence interval for the difference of two percentages.

For efficacy and safety endpoint(s), relative risks (i.e., the ratio of rates), confidence interval for the relative risks, the difference in rates and the confidence interval for difference in rates (using previously-described formulas), and p-values may also be presented for hypothesis generating purposes. The p-values will be based on either Pearson's Chi-square test or Fisher's exact test by checking the expected frequency for each cell in the 2x2 contingency table against Cochran's rule⁸, i.e., if the expected frequencies for all cells are ≥ 5 , then Pearson's Chi-square test will be used, otherwise Fisher's exact test will be used.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

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

1. 100(1- α) % Exact Clopper-Pearson Confidence Interval for A Single Proportion⁶

$$\text{Lower Confidence Limit} = \frac{x}{x + (n - x + 1)F_{1-\frac{\alpha}{2}}(2(n - x + 1), 2x)}$$

$$\text{Upper Confidence Limit} = \frac{(x + 1)F_{1-\frac{\alpha}{2}}(2(x + 1), 2(n - x))}{n - x + (x + 1)F_{1-\frac{\alpha}{2}}(2(x + 1), 2(n - x))}$$

where:

n = sample size

x = number of "events"

$F_{1-\frac{\alpha}{2}}(df_1, df_2)$ = the (1- α /2) F - statistic for degrees of freedom df_1 and df_2

2. 100(1- α) % Newcombe Score Confidence Interval for the Difference of Two Proportions⁷

a. 100(1- α) % Wilson Score Confidence Interval for A Single Proportion⁵

$$\text{Lower Confidence Limit} = \left(\hat{p} + Z_{\alpha/2}^2 / 2n - Z_{\alpha/2} \sqrt{(\hat{p}(1 - \hat{p}) + Z_{\alpha/2}^2 / 4n) / n} \right) / (1 + Z_{\alpha/2}^2 / n)$$

$$\text{Upper Confidence Limit} = \left(\hat{p} + Z_{\alpha/2}^2 / 2n + Z_{\alpha/2} \sqrt{(\hat{p}(1 - \hat{p}) + Z_{\alpha/2}^2 / 4n) / n} \right) / (1 + Z_{\alpha/2}^2 / n)$$

where:

$$\hat{p} = x / n$$

n = sample size

x = number of "events"

$Z_{\alpha/2}$ = 100(1- α /2)th percentile of the standard normal distribution

b. 100(1- α) % Newcombe Score Confidence Interval for the Difference of Two Proportions⁴

$$\text{Lower Confidence Limit} = (\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2} \sqrt{L_1(1-L_1)/n_1 + U_2(1-U_2)/n_2}$$

$$\text{Upper Confidence Limit} = (\hat{p}_1 - \hat{p}_2) + Z_{\alpha/2} \sqrt{U_1(1-U_1)/n_1 + L_2(1-L_2)/n_2}$$

where:

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

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

L_1 and U_1 are the lower and upper Wilson Score confidence limits for p_1

L_2 and U_2 are the lower and upper Wilson Score confidence limits for p_2

$Z_{\alpha/2}$ = 100(1 - $\alpha/2$)th percentile of the standard normal distribution

2.1.3 Propensity Score

Given that subjects in the two comparison groups (the 1-month DAPT arm vs XIENCE V USA historical control) are not randomized and thus may not have balanced baseline characteristics, the non-inferiority and superiority tests for 1-6 month period will be carried out through stratified analysis in the “1-month clear” population. The stratification will be performed through propensity scores (PS). For each individual a propensity score (i.e., predicted probability between 0 and 1) for group (1-month DAPT arm) membership will be calculated using logistics regression, with “group” as the outcome and baseline variables including demographic, lesion characteristics, and risk factors as the predictors. Subjects will be categorized into 5 groups based on the calculated propensity scores and their quintiles. Non-inferiority and superiority will then be carried out for the rate of the endpoint using the method described in sections below stratified by the propensity quintiles.

The propensity score modeling and design will be performed by an independent statistician who has no access to any outcome data of the XIENCE 28 Global study and the XIENCE V USA historical control for the integrity and interpretability of study results. The independent statistician will be blinded and have no access to clinical outcome and any follow-up information to avoid introducing bias into the analysis. [REDACTED]



2.1.3-1 Propensity Score Variable List

Based on prior clinical experience and clinical research, below is the list of variables that are considered to be related to assignments modeling building:

- Gender
- Age
- Creatinine
- Chronic anticoagulant
- History of stroke
- History of major bleeding
- Platelet
- Hb
- BMI
- Hypertension
- Dyslipidemia
- Prior PCI
- Prior CABG
- Prior MI
- Multivessel disease
- Clinical presentation (ACS (NSTEMI, ACS unstable angina) vs. non-ACS)
- Diabetes
- ACC/AHA lesion complexity

- Total lesion length per patient
- RVD
- Diameter stenosis%
- Bifurcation
- Number of lesions treated
- Number of vessels treated
- Number of stents per patient
- Total stent length per patient
- Discharge P2Y12
- Paris bleeding score
- PRECISE DAPT score



2.1.4 Hypothesis Testing

Primary endpoint analysis for 1-month DAPT

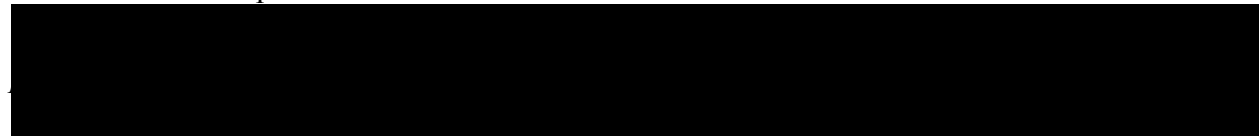
The XIENCE 28 Global study is powered based on primary endpoint of NACE between 1-month and 6-month follow-up. NACE is defined as the composite rate of all-cause death, all myocardial infarction (modified Academic Research Consortium [ARC]), stent thrombosis (ARC definite or probable), stroke or major bleeding (Bleeding defined by the Bleeding Academic Research Consortium [BARC] type 2-5) from 1 to 6 months. This primary endpoint will be evaluated between 1-month DAPT arm and XIENCE V USA historical control stratified by propensity score quintiles in the primary analysis population. The null (H_0) and alternative (H_A) hypothesis are of the following form:

$$H_0: NACE_{1m-DAPT} - NACE_{Xience\ V\ USA} \geq \delta$$

$$H_A: NACE_{1m-DAPT} - NACE_{Xience\ V\ USA} < \delta$$

Where δ is the non-inferiority margin. The test will be carried out with a one-sided significance level of 0.05 and a non-inferiority margin (δ) of 4.0%.

The stratified Farrington-Manning method will be performed to test non-inferiority (NI) of 1-month DAPT compared to standard DAPT duration from XIENCE V USA historical control.



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

2.1.5 Survival Analyses

Survival analysis may be conducted to analyze time-to-event variables. Subjects without events will be censored at their last known event-free time point. Survival curves will be constructed using Kaplan-Meier estimates.

Summary tables for the endpoints will include failure rates (Kaplan-Meier estimates). For the primary analysis report, all available data will be used.

2.2 Endpoint Analyses



2.2.1 Primary Endpoint Analysis

Primary Analysis:

A non-inferiority test will be performed on the primary endpoint between 1-month and 6-month follow up for the primary analysis population of the 1-month DAPT arm and the XIENCE V USA historical control stratified by propensity score quintiles as described in section 2.1.4.

To ensure all subjects to be included in this analysis, multiple imputations will be performed in calculating the propensity scores.

Secondary Analysis (as a sensitivity analysis):

In addition to the primary analysis stratified by propensity score quintile, a non-inferiority test will be performed on the primary endpoint stratified by propensity score quartile, with the same methodology as described above.

2.2.3 Secondary Endpoint Analyses

The secondary clinical endpoints will be descriptively analyzed for both the primary analysis population and all registered subjects without propensity stratification.

2.3 Subgroups for Analysis

All of the following subgroup analyses are intended for the primary analysis population. The comparison between the 1-month DAPT arm and the XIENCE V USA historical control will be analyzed descriptively within each quintile for the primary in a specific subgroup. The above quintiles are based on the overall PS, not PS built within each subgroup, as baseline characteristics of subjects are likely to be comparable in each quintile of the overall PS.

2.3.1 Sex

Sex-specific subgroup analyses will be performed on the primary analysis population for the primary endpoint stratified by the overall PS.

2.3.2 Diabetes

Diabetic subgroup analysis will be performed on the primary analysis population for the primary endpoint stratified by the overall PS. Analyses will be performed within the following subgroups:

- All diabetes mellitus, defined as any diabetics with or without medical treatment
- Non diabetes mellitus.

2.3.3 Other Subgroups

The following subgroups analyses will be evaluated for the primary analysis population for the primary endpoint stratified by the overall PS. Analyses will be performed within the following subgroups:

- Age (age \geq median vs $<$ median)
- Clinical presentation (ACS NSTEMI, ACS unstable angina, non-ACS)

2.4 Analysis Window

- 6 months
- 12 months

2.5 Handling of Missing Data

The primary analysis will be evaluated after propensity score stratification. To handle missing data in propensity score building, multiple imputation methods will be performed for baseline characteristics to compute propensity scores from these datasets.

All other analyses will be based on available data with missing data excluded. Any unused or spurious data will be noted as appropriate in the final report.

2.6.1 Multiple Center Effect

Analysis will be performed by pooling data across study sites.

The 1-month DAPT arm will have 50 sites globally. Subject registration is capped at 120 per site. This cap per site will prevent the scenario where the results from a few sites dominate the overall study result. For the analysis of center effect, data from smaller sites may be combined for the analysis. Smaller sites are defined as sites with fewer than 20 subjects per site.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

2.7 Adjustments for Covariates

Unless otherwise specified, no adjustments for covariates will be made for any of the variables in the analyses.

2.8 Multiplicity Issues

No multiplicity adjustment is necessary because there is only one test on the primary endpoint

2.9 Sensitivity Analysis

The primary endpoint between 1-month and 6-month follow-up will be analyzed descriptively for the 1-month clear population removing the patients who do not have antiplatelet medication compliance after 1 month.

For the 1-month DAPT arm, antiplatelet medication non-compliance beyond 1-month follow-up for 1-month clear population is defined as patients who resume P2Y12 inhibitor for more than 7

consecutive days, and/or interrupt aspirin for more than 7 consecutive days between 1-month and 12-month follow up.

[REDACTED]

- [REDACTED]
- [REDACTED]

[REDACTED]

[REDACTED]

In addition, a sensitivity analysis will be performed for the primary endpoint analysis to evaluate the impact of the missing outcome. The analysis will be carried out based on the primary analysis population, and by imputing the missing outcomes for each imputed baseline PS dataset, and Rubin's combination rule ^[29] will be used to consolidate the final analysis for the 10 duplicates of the imputed dataset.

2.10 Documentation and Other Considerations

All analyses will be performed using SAS[®] for Windows, version 9.1 or higher.

3. ACRONYMS AND ABBREVIATIONS

Acronym or Abbreviation	Complete Phrase or Definition
ACS	Acute Coronary Syndrome
AMI	Acute Myocardial Infarction
ARC	Academic Research Consortium
BARC	Bleeding Academic Research Consortium
BMI	Body Mass Index
CABG	Coronary artery bypass grafting
CTO	Chronic Total Occlusion
DAPT	Dual Antiplatelet Therapy
DMR	Death/MI/Revascularization
EECSS	Everolimus Eluting Coronary Stent System
Hb	Hemoglobin
HBR	High Bleeding Risk
LVEF	Left Ventricular Ejection Fraction
MI	Myocardial Infarction
NACE	Net Adverse Clinical Endpoint
NSTEMI	non ST-segment elevation MI
PCI	Percutaneous Coronary Intervention
PS	Propensity Score
RVD	Reference Vessel Diameter
SAP	Statistical Analysis Plan
ST	Stent Thrombosis
STEMI	ST-segment elevation myocardial infarction
TLF	Target Lesion Failure
TLR	Target Lesion Revascularization
TVR	Target Vessel Revascularization
US	United States

4. REFERENCES

1. Li, H., Mukhi, V., Lu, N., Xu, Y.L. and Yue, L.Q., 2016. A Note on Good Practice of Objective Propensity Score Design for Premarket Nonrandomized Medical Device Studies with an Example. *Statistics in Biopharmaceutical Research*, 8(3), pp.282-286.
2. D'Agostino, R.B., 1998. Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*, 17(19), pp.2265-2281.
3. Farrington, C. P., Manning, G., Test Statistics and Sample Size Formulae for Comparative Binomial Trials with Null Hypothesis of Non-Zero Risk Difference or Non-Unity Relative Risk. *Statistics in Medicine*, 1990, 9, 1447-1454.
4. Hintze J. NCSS and PASS. Number Cruncher Statistical Systems. Kaysville, Utah. 2005.
5. SAS/STAT® 9.2 User's Guide, Second Edition, SAS Institute Inc., Cary, NC, 2009.
6. Clopper C. J., Pearson E. S., The Use of the Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika*, 1934, 26, 404-413.
7. Newcombe, R. G., Interval estimation for the difference between independent proportions: comparison of eleven methods, *Statistics in Medicine*, 1998, 17, 873-890.
8. Cochran, W. G., Some Methods for Strengthening the Common χ^2 Test. *Biometrics*, 1954, 10, 417-451.
9. Wilson, E. B., Probable inference, the law of succession, and statistical inference, *Journal of the American Statistical Association*, 1927, 22, 209-212.
10. Draft Guidance for Industry and Food and Drug Administration Staff: Evaluation of Sex Differences in Medical Device Clinical Studies, December 19, 2011, CDRH
11. Gail, M., Simon, R., Testing for Qualitative Interactions between Treatment Effects and Patient Subsets, *Biometrics*, 1985, 41, 361-372.
12. Dmitrienko, A. et. al., *Analysis of Clinical Trials Using SAS: A Practical Guide*, SAS Institute, Cary, NC, 2005.
13. Fleming TR& Harrington DP 1991. *Counting Processes and Survival Analysis*. New York: Wiley.
14. Miller R & Halpern J 1982. Regression with censored data. *Biometrika* 69, 521-31.
15. Allison PD. *Survival Analysis Using the SAS system: A Practical Guide*, Cary, NC: SAS Institute Inc., 1995.
16. Kaplan EL, Meier P. Non-parametric estimation from incomplete observations. *J Am Stat Assoc* 1958; 53:457-481.
17. Therneau TM, Grambsch PM. *Modeling survival data: extending the Cox model*. New York: Springer-Verlag; 2000.
18. Allison PD. Multiple Imputation for Missing Data. *Sociological Methods & Research*, Feb 2000; Vol. 28 No. 3:301-309.

19. Barnes SA, Lindborg SR, Seaman JW. Multiple Imputation Techniques in Small Sample Clinical Trials. *Stat Med* 2006 Jan 30; 25(2):233-45.
20. Clopper CJ, Pearson E. The Use of the Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika* 1934; 26:404-413.
21. Lavori PW, Dawson R, Shera D. A Multiple Imputation Strategy for Clinical Trials with Truncation of Patient Date. *Statistics in Medicine* 1995; 14:1913-1925.
22. Li X, Mehrotra DV, Barnard J. Analysis of Incomplete Longitudinal Binary Data Using Multiple Imputations. *Stat Med.* 2006 Jun 30; 25(12):2107-24.
23. Lavori PW, Dawson R, Shera D. A multiple imputation strategy for clinical trials with truncation of patient data. *Stat Med.* 1995; 14:1913–1925.
24. Little RJA, Rubin D B. *Statistical Analysis with Missing Data*. New York: John Wiley & Sons; 1987.
25. Rubin DB. Inference and missing data. *Biometrika.* 1976; 63:581-592.
26. Little RJA, Schenker N. *Handbook of Statistical Methodology-Missing Data*. New York: Plenum Press; 1995.
27. D’Agostino, R. B., Jr., Rubin, D. B. (2000). Estimating and using propensity scores with partially missing data. *Journal of the American Statistical Association* 95:749–759
28. Yue, L.Q., Lu, N. and Xu, Y., 2014. Designing premarket observational comparative studies using existing data as controls: challenges and opportunities. *Journal of biopharmaceutical statistics*, 24(5), pp.994-1010.
29. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Wiley: New York, 1987.
30. Nelson Lu, Yunling Xu, and Lilly Q. Yue (2020) Some Considerations on Design and Analysis Plan on a Nonrandomized Comparative Study Using Propensity Score Methodology for Medical Device Premarket Evaluation, *Statistics in Biopharmaceutical Research*, 12:2, 155-163, DOI: 10.1080/19466315.2019.1647873

[REDACTED]

[REDACTED]

[REDACTED]

		[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]