Janssen Research & Development *

Statistical Analysis Plan

Medically Ill Patient Assessment of Rivaroxaban Versus Placebo IN Reducing Post-Discharge Venous Thrombo-Embolism Risk (MARINER)

Protocol RIVAROXDVT3002; Phase 3

JNJ-39039039; BAY 59-7939 (rivaroxaban)

Amendment 1

Status: Approved
Date: 4 April 2017
Prepared by: Janssen Research & Development, LLC
Document No.: EDMS-ERI-82523399:5.0

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

TABLE OF CONTENTS **2**

LIST OF IN-TEXT TABLES **3**

LIST OF IN-TEXT FIGURES **3**

AMENDMENT HISTORY **4**

ABBREVIATIONS **9**

1. INTRODUCTION **10**
1.1. Trial Objectives 10
1.2. Trial Design 12
1.3. Statistical Hypotheses for Trial Objectives 14
1.4. Sample Size Justification 15
1.5. Randomization and Blinding 16

2. GENERAL ANALYSIS DEFINITIONS **16**
2.1. Definition of Trial Dates 16
2.2. Pooling Algorithm 18
2.3. Analysis set and Analysis Phase 18
2.3.1. Primary Efficacy Analysis Set 19
2.3.2. Secondary Efficacy Analysis Set 19
2.3.3. Safety Analysis Set 19
2.4. Definition of Subgroups 19

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW **20**

4. SUBJECT INFORMATION **21**
4.1. Demographics and Baseline Characteristics 21
4.2. Disposition Information 22
4.3. Treatment Compliance 23
4.4. Extent of Exposure 23
4.5. Protocol Deviations 23
4.6. Prior and Concomitant Medications 24

5. EFFICACY **24**
5.1. Analysis Specifications 24
5.1.1. Level of Significance 24
5.1.2. Data Handling Rules 24
5.2. Primary Efficacy Outcome 24
5.2.1. Definition 24
5.2.2. Analysis Methods 25
5.3. Secondary Efficacy Outcomes 27
5.3.1. Definition 27
5.3.2. Analysis Methods 27
5.4. Multiple Testing Procedure 27
5.5. Other Efficacy Outcomes 28
5.5.1. Definition 28
5.5.2. Analysis Methods 29

6. SAFETY **29**
6.1. Bleeding Events 29
6.2. Adverse Events 30
6.3. Clinical Laboratory Tests 30

6.4. Vital Signs and Physical Examination Findings 30

7. BENEFIT/RISK ASSESSMENT 31

8. PHARMACOKINETICS 33

9. HEALTH ECONOMICS 34

REFERENCES 35

ATTACHMENTS 36
Attachment 1: Template of CONSORT Statement 2010 flow diagram 36

LIST OF IN-TEXT TABLES

Table 1: Approximate Number of Subjects Needed to Observe 161 Events 16
Table 2: Conditional Power (CP) Based on the Assumption of the Various Underlying RRRs in the Remaining Study 21

LIST OF IN-TEXT FIGURES

Figure 1: Schematic Overview of the Study Design by Country 14

AMENDMENT HISTORY

SAP AMENDMENT 1

The original Statistical Analysis Plan (SAP) was finalized and issued on May 29th, 2014, based on the Clinical Protocol RIVAROXDVT3002 Amendment INT-2, finalized on April 11th, 2014.

This document is the SAP Amendment 1 that incorporates further changes and clarifications based on the Protocol Amendments INT-3, INT-4, INT-5, INT-6, and the Protocol Amendment INT-7 that was finalized on March 31th, 2017.

Listed below are further changes and clarifications in SAP Amendment 1 (with major changes underlined and bolded).

Topic or Section	Description of further changes or clarification in SAP Amendment 1
Section 1 Introduction	Clarification on SAP Amendment 1: This SAP Amendment 1 is based on the Protocol RIVAROXDVT3002 Amendment INT-3, INT-4, INT-5, INT-6 and the Protocol RIVAROXDVT3002 Amendment INT-7, finalized on March 31th, 2017. Unless the key analysis method for efficacy and/or safety needs to be updated, a SAP amendment will not take place, regardless of a protocol amendment.
Section 1.1 Trial Objectives	Clarification of terminology and **added 1 endpoint in secondary efficacy endpoints and deleted 1 exploratory endpoint**: Added: The composite of symptomatic VTE (lower extremity DVT and non-fatal PE), myocardial infarction (MI), non-hemorrhagic stroke and cardiovascular (CV) death (death due to a known CV cause and death in which a CV cause cannot be ruled out; by this definition, a VTE-related death is considered a CV death) Deleted: The composite of the primary efficacy outcome and cardiovascular (CV) events (fatal and non-fatal myocardial infarction [MI] and non-hemorrhagic stroke)
Section 1.2 Trial Design	Clarification on the randomization, the first dose date, visit schedules, interim analysis (IA) timing.

	Clarification: The total number of subjects that may be randomized was increased from approximately 9,000 to approximately 12,000 subjects Added a statement: The total number of VTE-related death events may also be taken into account when deciding to stop randomization. **Clarification: when approximately 50% of subjects with an adjudicated primary efficacy outcome have been observed (about 50% of the expected total number of events based on the actual observed blinded pooled event rate even if this event rate is lower than the assumed rate)**
Section 1.3 Statistical Hypotheses for Trial Objective	**Added 1 endpoint in secondary efficacy endpoints**: The composite of symptomatic VTE (lower extremity DVT and non-fatal PE), myocardial infarction (MI), non-hemorrhagic stroke and cardiovascular (CV) death (death due to a known CV cause and death in which a CV cause cannot be ruled out; by this definition, a VTE-related death is considered a CV death)
Section 1.4 Sample size Justification	**Clarification: The total number of subjects that may be randomized was increased from approximately 9,000 to approximately 12,000 subjects** Added a statement: The total number of VTE-related death events may also be taken into account when deciding to stop randomization Added more cases in Table 1:Approximate Number of Subjects Needed to Observe 161 Events
Section 1.5 Randomization and Blinding	Clarification: changed CrCl 30 to 49 to CrCl ≥30 and <50
Section 2.1 Definition of Trial Dates	Changed 'the date of the last visit at which study drug is dispensed plus the number of pills in the drug bottle(s) dispensed at that visit' to 'the latest logically possible date on or after the first dose date'

Section 2.3 Analysis Sets and Analysis Phase	Changed 'Major' protocol deviation to 'Key' protocol deviation Clarification: updated inclusion/exclusion numbers
Section 2.3.1 Primary Efficacy Analysis Set	Added: Subjects will be analyzed according to the treatment randomized.
Section 2.4 Definition of Subgroups	Clarification and added 'and not limited to the following subgroups', Use of US approved thromboprophylaxis (ie UFH, enoxaparin and dalteparin) vs other and Baseline D-dimer level (> 2xULN vs ≤ 2xULN)
Section 3 Interim Analysis and Data Monitoring Committee	**Clarification: An IA for futility will be conducted when approximately 50% of subjects with adjudicated primary efficacy outcome have been observed (about 50% of the expected total number of events based on the actual observed blinded pooled event rate even if this event rate is lower than the assumed rate)** Observed primary efficacy outcome incidence rate by Day 45 in placebo at IA is too low that findings are not likely clinically meaningful
Section 4.1 Demographics and Baseline Characteristics	Clarification: changed congestive heart failure to heart failure (HF) with reduced ejection fraction Added 'and not limited to the following demographics and baseline characteristics', Use of US approved thromboprophylaxis (ie UFH, enoxaparin and dalteparin) vs other and Baseline D-dimer level (> 2xULN vs ≤ 2xULN)
Section 5.2.1 Definition	Clarification: (death due to PE or death in which PE cannot be ruled out as the cause)
Section 5.2.2 Analysis Methods	Added: The additional sensitivity analyses of the primary efficacy outcome may be necessary to exclude subjects (e.g., significant quality issues, Good Clinical Practice (GCP) violations, data integrity issues, regulatory agency commitment)

Section 5.3.1 Definition	Added: The composite of symptomatic VTE (lower extremity DVT and non-fatal PE), myocardial infarction (MI), non-hemorrhagic stroke and cardiovascular (CV) death (death due to a known CV cause and death in which a CV cause cannot be ruled out; by this definition, a VTE-related death is considered a CV death)
Section 5.4 Multiple Testing Procedure	**Added: The composite of symptomatic VTE, myocardial infarction (MI), non-hemorrhagic stroke and cardiovascular (CV) death (death due to a known CV cause and death in which a CV cause cannot be ruled out; by this definition, a VTE-related death is considered a CV death)**
Section 5.5.1 Definition	Deleted: The composite of the primary efficacy outcome and cardiovascular (CV) events (fatal and non-fatal myocardial infarction [MI] and non-hemorrhagic stroke) Added: Fatal MI or fatal stroke are defined as death occurred within 30 days of the MI or stroke and death was directly related to the MI/stroke respectively
Section 6.1	Added: bleeding by Bleeding Academic Research Consortium (BARC) classification
Section7 Net Clinical Benefit	Clarification: (lower extremity DVT, non-fatal PE), (death due to PE or death in which PE cannot be ruled out as the cause) Added 3 risk endpoints: Non-fatal, non-critical site bleeding requiring transfusion ≥2 units, Non-fatal, non-critical site Hgb drop ≥ 2 g/dL, and Non-fatal, non-critical site Hgb drop ≥ 2 g/dL not associated with transfusion ≥2 units Added: A bivariate analysis may be explored as post hoc analysis

Section 8 Pharmacokinetics	Added: Statistical summary of concentration (e.g. geometric mean, standard deviation, median, 5th and 95th percentile, range) by nominal sampling time window as specified in protocol will be provided. Statistical summary will also be stratified by CrCl groups. Results will be provided separately from the clinical study report

ABBREVIATIONS

ACM	all-cause mortality
AE	adverse event
ATC	Anatomical Therapeutic Chemical
BARC	Bleeding Academic Research Consortium
CCU	cardiac care unit
CEC	Clinical Event Committee
CI	confidence interval
CMH	Cochran–Mantel–Haenszel
CrCl	creatinine clearance
DPS	Data Presentation Specifications
DVT	deep vein thrombosis
eCRF	electronic case report form
EC	Executive Committee
EOS	end of study
EOT	end of treatment
ER	emergency room
GCP	Good Clinical Practice
HF	heart failure
HR	hazard ratio
ICH	International Conference on Harmonization
ICU	intensive care unit
IDMC	Independent Data Monitoring Committee
IMPROVE	International Medical Prevention Registry On Venous Thromboembolism
ISTH	International Society on Thrombosis and Haemostasis
ITT	intention-to-treat
LMWH	low molecular weight heparin
LOS	length of stay
MedDRA	Medical Dictionary for Regulatory Activities
PE	pulmonary embolism
PK	pharmacokinetic(s)
PP	per-protocol
RRR	relative risk reduction
SAE	serious adverse event
SAP	Statistical Analysis Plan
TE	treatment-emergent
UFH	unfractioned heparin
VTE	venous thromboembolism

1. INTRODUCTION

This Statistical Analysis Plan (SAP) Amendment 1 specifies definitions of analysis sets, key derived variables, and statistical methods for analysis of efficacy and safety for the Phase 3 study RIVAROXDVT3002 (also known as MARINER). This SAP Amendment 1 is based on the Protocol RIVAROXDVT3002 Amendment INT-3, INT-4, INT-5, INT-6 and the Protocol RIVAROXDVT3002 Amendment INT-7, finalized on March 31th, 2017. Unless the key analysis method for efficacy and/or safety needs to be updated, a SAP amendment will not take place, regardless of a protocol amendment. Titles, mock-ups and programming instructions for all statistical outputs (tables, figures, and listings) are provided in a separate document entitled Data Presentation Specifications (DPS).

An Independent Data Monitoring Committee (IDMC) will monitor safety during this study, review interim safety/efficacy results, and make a recommendation whether the study should be continued as planned, modified or terminated early for futility or safety. There is no plan to stop the study for the superiority of the primary efficacy outcome. Safety data to be reviewed by the IDMC periodically and additional details for the planned interim analysis are specified in the IDMC SAP, which is specifically prepared for IDMC reviews.

1.1. Trial Objectives

The primary objective is to assess the efficacy and safety of rivaroxaban, compared with placebo

in the prevention of symptomatic venous thromboembolism [VTE] (lower extremity deep vein thrombosis [DVT] and non-fatal pulmonary embolism [PE]) and VTE-related death (death due to PE or death in which PE cannot be ruled out as the cause) post-hospital discharge in high-risk, medically ill patients.

The secondary objectives are to compare rivaroxaban with placebo in the following post-hospital discharge outcomes in high-risk, medically ill patients:

- VTE-related death (death due to PE or death in which PE cannot be ruled out as the cause)
- Symptomatic VTE (lower extremity DVT and non-fatal PE)
- The composite of symptomatic VTE (lower extremity DVT and non-fatal PE) and all-cause mortality (ACM)

- The composite of symptomatic VTE (lower extremity DVT and non-fatal PE), myocardial infarction (MI), non-hemorrhagic stroke and cardiovascular (CV) death (death due to a known CV cause and death in which a CV cause cannot be ruled out; by this definition, a VTE-related death is considered a CV death)
- All-cause mortality (ACM)

The exploratory objectives are to compare rivaroxaban with placebo, in the following post-hospital discharge outcomes in high-risk, medically ill patients:

- Symptomatic lower extremity DVT
- Symptomatic non-fatal PE
- Symptomatic upper extremity DVT
- MI
- Non-hemorrhagic stroke
- Re-hospitalization for symptomatic VTE (lower extremity DVT and non-fatal PE) within 30 days after randomization

The pharmacokinetic (PK) objective is to assess the kinetics of rivaroxaban in high-risk, medically ill patients, to describe drug exposure based on creatinine clearance (CrCl).

The Health Economics objective is to assess the following cost drivers, ie, re-hospitalization (including length of stay [LOS]), emergency room (ER) visits, intensive care unit (ICU) and cardiac care unit (CCU) stays, including LOS and subject discharge disposition between the 2 treatment groups.

The safety objectives are to compare rivaroxaban with placebo in the following bleeding outcomes in high-risk, medically ill patients:

- Major bleeding using validated International Society on Thrombosis and Haemostasis (ISTH) bleeding criteria.
- Non-major clinically relevant bleeding
- Other bleeding

Overall safety will also be assessed.

1.2. Trial Design

This is a multicenter, prospective, randomized, double-blind, placebo-controlled, event-driven study designed to evaluate rivaroxaban, compared with placebo, in the prevention of symptomatic VTE (lower extremity DVT and non-fatal PE) events and VTE-related deaths (death due to PE or death in which PE cannot be ruled out as the cause) for a period of 45 days post-hospital discharge. The first dose of study drug should be administered no later than the day after the subject leaves the hospital and as soon after randomization (Day 1) as possible. Subjects will be instructed to discontinue study drug after they take a dose on Day 45. The subject population comprises men and women age 40 and over who had been hospitalized for at least 3 and no more than 10 consecutive days prior to randomization for a specific acute medical illness and have other risk factors for VTE. Subjects who meet all of the inclusion and none of the exclusion criteria will be randomly assigned to receive rivaroxaban or placebo. Randomization should occur on the same day as or the day after the subject leaves the hospital, and may occur at the hospital, clinic, or other discharge destination. Subjects will be randomized by strata (by subjects with a CrCl ≥30 mL/min and <50 mL/min versus subjects with CrCl ≥50 mL/min, and by country) by an interactive web response system (IWRS). After randomization, subjects will receive double-blind treatment with rivaroxaban 10 mg (7.5 mg for subjects with a CrCl ≥30 mL/min and <50 mL/min) daily or matching placebo daily.

The targeted total number of primary efficacy outcome in this event-driven study is 161. It is estimated that a total of approximately 8,000 subjects will need to be randomized to either rivaroxaban or placebo in 1:1 ratio. This estimate is based on an estimated placebo incidence rate of the primary efficacy outcome of 2.5%. In the event that the actual observed blinded pooled event rate is lower than the assumed rate, more subjects may be enrolled to accumulate additional outcome events. Randomization in this study may be stopped at approximately 12,000 subjects for administrative considerations even if the targeted 161 events have not been observed by then. The total number of VTE-related death events may also be taken into account when deciding to stop randomization.

An interim analysis for futility will be conducted when approximately 50% of subjects with an adjudicated primary efficacy outcome have been observed (about 50% of the expected total number of events based on the actual observed blinded pooled event rate even if this event rate is lower than the assumed rate).

Prior to randomization, all subjects must have received treatment during the index hospitalization with LMWH or UFH. The study consists of a screening phase, a 45-day double-blind treatment phase, and a 30-day follow-up. The total duration for a subject who completes the study after randomization is expected to be 75 days.

If the subject permanently discontinues treatment before Day 45, he/she should be instructed to complete an unscheduled visit and the remaining scheduled visits, including the Day 7 and Day 21 visits (if not yet completed) and the Day 45 (+4d) and Day 75 (±5d) visits

If a subject withdraws consent from the study or is lost to follow-up, his or her vital status and outcomes will be collected at the day 75 (±5d) visit either by telephone or in person, or if applicable, by a review of subject's medical or public records unless this contact is not allowed by local regulations.

For population PK analyses, based on the model performance and high variability of the PK parameters observed in the MAGELLaN study, at the Day 7 visit 2 blood samples (at pre-dose; 1-4 hours after study drug administration) and again at Day 21 visit (Visit 4) 2 blood samples (at 3 to 7 and 7 to 12 hours after study drug administration) will be collected for each subject for about 600 subjects (300 rivaroxaban subjects [200 subjects in the 10 mg QD group and 100 subjects in the 7.5 mg QD group] and 300 in the placebo group). The sparse blood samples for PK analyses will be collected in selected sites only.

An Executive Committee (EC), an Independent Data Monitoring Committee (IDMC) and a Clinical Event Committee (CEC) will be commissioned for this study.

The EC has overall responsibility for the design, conduct and reporting of the study. The EC will monitor overall safety during the study and will receive any recommendations from the IDMC regarding possible additional analyses or modifications to the study and decide whether to accept them. The EC will oversee the implementation of any modifications to the study.

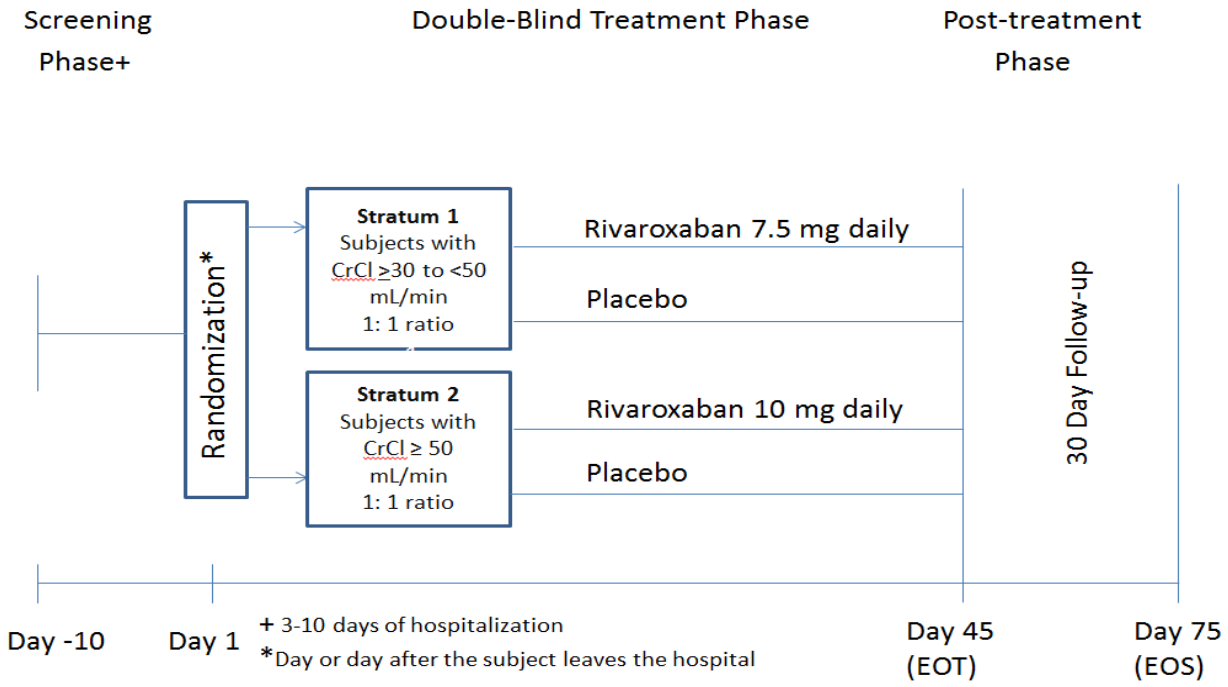
The IDMC will review unblinded safety data periodically to ensure the safety of study subjects. If necessary or requested by the IDMC, subject level unblinded data may be provided to the IDMC. In addition, the IDMC will review results of the planned interim analysis and make a recommendation whether the study should be terminated prematurely due to futility. Detailed stopping guidelines are specified in the section 3.

The Clinical Event Committee (CEC) will adjudicate all suspected outcome events in a consistent and unbiased manner according to the definitions in the CEC charter while blinded to treatment assignment.

All committees will be governed by separate charters.

A diagram of the study design is provided below in Figure 1:

Figure 1: Schematic Overview of the Study Design by Country



1.3. Statistical Hypotheses for Trial Objectives

The primary statistical alternative hypothesis is that rivaroxaban is superior to placebo on the primary efficacy outcome, that is, time to event for the rivaroxaban group is stochastically later than that for the placebo group from randomization up to Day 45. More specifically, the survival function for the placebo group is lower than that for the rivaroxaban group. The null hypothesis is the negation of the alternative hypothesis, that is, the survival function for the placebo group is not lower than that for the rivaroxaban group. As a further illustration of the hypotheses, for the case where hazard functions are proportional, the above alternative hypothesis can be expressed as that the hazard ratio (HR) of rivaroxaban versus placebo is less than 1, and the null hypothesis is that the HR is at least 1.

The secondary hypotheses are that rivaroxaban is superior to placebo in the prevention of 1) VTE-related death (death due to PE or death in which PE cannot be ruled out as the cause), 2) symptomatic VTE (lower extremity DVT and non-fatal PE), 3) the composite of symptomatic VTE and ACM, 4) the composite of symptomatic VTE (lower extremity DVT and non-fatal PE), MI, non-hemorrhagic stroke and CV death (death due to a known CV cause and death in which a CV cause cannot be ruled out; by this definition, a VTE-related death is considered a CV death) and 5) ACM alone.

1.4. Sample Size Justification

This is an event driven study. The targeted total number of primary efficacy outcome is 161, based on the ITT analysis set and Up-to-Day-45 analysis phase. If a subject has multiple events, only the first is counted towards study size determination.

This targeted total number of events is determined using statistical software East 5.3 based on the primary efficacy analysis (defined later) and the following assumptions:

- 40% relative risk reduction (RRR) in the primary efficacy outcome based on the ITT analysis set and Up-to-Day-45 analysis phase (RRR is defined as 1 minus the hazard ratio (HR) of rivaroxaban versus placebo.)
- Power of 90% assuming the above RRR
- Two-sided significance level of 0.05.

To observe the targeted 161 events, it is estimated that a total of approximately 8,000 subjects will need to be randomized to either rivaroxaban or placebo in 1:1 ratio. This estimate is based on an estimated placebo incidence rate of the primary efficacy outcome of 2.5%. In the event that the actual observed blinded pooled event rate is lower than the assumed rate, more subjects may be enrolled to accumulate additional outcome events. Randomization in this study may be stopped at approximately 12,000 subjects for administrative considerations even if the targeted 161 events have not been observed by then. The total number of VTE-related death events may also be taken into account when deciding to stop randomization.

Table 1 below additionally shows the approximate numbers of randomized subjects needed to observe 161 events based on several other placebo incidence rates.

Table 1: Approximate Number of Subjects Needed to Observe 161 Events

Placebo Incidence Rate	Number of Subjects*
3.0%	6,700
2.75%	7,300
2.5%	8,000
2.25%	9,000
2.0%	10,100
1.5%	13,400
1.25%	16,100
*calculated based on 40% RRR	

1.5. Randomization and Blinding

Central randomization will be implemented in this study. Subjects will be randomly assigned to 1 of 2 treatment groups based on a computer-generated randomization schedule prepared before the study under the supervision of the sponsor. The randomization will be balanced by using randomly permuted blocks and will be stratified by subjects with CrCl ≥30 and <50 mL/min versus subjects with CrCl ≥ 50 ml/min and by country.

2. GENERAL ANALYSIS DEFINITIONS

Summaries by treatment group using appropriate descriptive statistics will be provided for all study variables including demographic and baseline characteristics. No imputation will be applied, unless specified otherwise in the SAP. Descriptive statistics such as mean, median, standard deviation, minimum, and maximum will be used to summarize continuous variables. Counts and percentages will be used to summarize categorical variables. Kaplan-Meier method will be used to summarize time-to-event variables. Graphical data displays may also be used to summarize the data.

Unless stated otherwise, subjects will be analyzed according to the treatment group they are assigned, irrespective of the actual treatment received.

Unless stated otherwise, all adjudicated efficacy and bleeding outcomes will be used in the analyses.

2.1. Definition of Trial Dates

Trial reference start date: the date of randomization of the subject.

Trial reference end date: the date of the last trial-related procedure for the subject and is defined as the maximum of the following dates:

- Dates of all study-related visits (including scheduled or unscheduled visits)
- Dates of all study-related procedures, findings and events, including, but not limited to, adverse events, concomitant medications, disposition, clinical laboratories, vital signs, death

First dose date: the date on which the first dose of study drug is taken by the subject. If missing or incomplete for a subject who takes study drug, the first dose date is set to the earliest logically possible date on or after randomization. More specifically, the first dose date is defined as the maximum of the randomization date and the first day of the month if only day is missing, and is defined as the randomization date if day and month are missing.

Last dose date: the date on which the last dose of study drug is taken by the subject. If missing or incomplete for a subject who takes study drug, the last dose date is in general set to the latest logically possible date on or after the first dose date.

To ensure consistency with other data of the trial, more specifically, the above date will be further

- capped by (that is, taking the minimum with) Day 45 date, death date, trial reference end date, and the upper bound of the logically possible range if partial date is available, and
- further bounded below by (that is, taking the maximum with) the (imputed) first dose date, any complete temporary drug stop and re-start date, or the lower bound of the logically possible range.

Upper bound of the logically possible range is defined as the last day of the month if only day is missing and is defined as the last day of the following month after randomization if day and month are missing. Lower bound of the logically possible range is defined the maximum of the randomization date and the first day of the month if only day is missing and is defined as the randomization date if day and month are missing.

Last outcome-evaluation date: the date of last study visit (clinic/phone) while the subject is alive during which the outcome status is ascertained.

2.2. Pooling Algorithm

No pooling by country will be done for by-country analysis.

2.3. Analysis set and Analysis Phase

Each analysis involves the following 2 aspects: 1) analysis set, specifying the subjects to be included; and 2) analysis phase, specifying the time window within which data will be included. Key analysis sets and analysis phases are defined below.

Analysis sets

- Intention-to-Treat (ITT) (as termed Full Analysis Set in the International Conference on Harmonisation [ICH] E9 guideline): This analysis set consists of all randomized subjects who have a signed valid informed consent.
- Per-Protocol (PP): This analysis set is a subset of the ITT analysis set.

Subjects with the key protocol deviations will be excluded from the PP analysis set. The categories and terms of the key protocol deviations will include the following:

- Did not meet the following inclusion or met the following exclusion criteria;
  - INT-2, INT-3, INT-4: subject with index hospitalization > 14 days, inclusion criteria 1, 2, 4, 6, exclusion criteria 10, 11, 14, 22, 24, 27
  - INT-6, INT-7: inclusion criteria 1.2, 2.2, 4.3, 6.1, exclusion criteria 10.2, 11.1, 14.1, 22.2, 24.2 without c
- Did not take any assigned study drug or took wrong study drug
- Did not discontinue study drug permanently according to the protocol (The subject becomes pregnant, develops any condition which requires permanent anticoagulation, has a fall in CrCl to below 20 mL/min during the study, or has a hemorrhagic stroke or intracranial bleeding)
- Had been taking prohibited concomitant therapies as specified in the protocol
  - INT-2: 2, 3,
  - INT-3, INT-4: 2, 3, 4, 6, 7, 8
  - INT-6, INT-7: 2.2, 3.2, 4.1, 7.1 without c and i

- Safety: This analysis set is a subset of the ITT analysis set, consisting of subjects who receive at least one dose of study drug.

Analysis phases

- Up-to-Day-45: This analysis phase includes all data from randomization to Day 45 (inclusive).

For time to event analyses, subjects who do not have events to be analyzed on or before Day 45 will be censored on Day 45 or the last outcome-evaluation date whichever occurs first.

- On-Treatment: This analysis phase includes all data from randomization to 2 days after the last dose of the study drug (inclusive).

For time to event analyses, subjects who do not have events to be analyzed during this phase will be censored on the last dose + 2 or the last outcome-evaluation date whichever occurs first.

2.3.1. Primary Efficacy Analysis Set

The primary efficacy analysis set for all efficacy outcomes is the ITT analysis set. Subjects will be analyzed according to the treatment randomized.

2.3.2. Secondary Efficacy Analysis Set

The PP analysis set will be used to assess the consistency with the primary efficacy analysis.

2.3.3. Safety Analysis Set

The primary safety analysis set for all safety outcomes is the safety analysis set. Subjects will be analyzed according to the actual treatment received.

2.4. Definition of Subgroups

Homogeneity of treatment effects, both in HR and direction, will be assessed in the following subgroups classified at baseline and not limited to the following subgroups:

- Age (< 65 vs ≥65; <75 vs ≥75 years)
- Sex (men vs women)
- Race (White vs others)
- Race (White, Black, Asian, others)

- Geographic region (North America, South America, Western Europe, Eastern Europe, other)
- CrCl (≥30 to < 50, ≥50 to <80, ≥80 mL/min)
- Body Mass Index (<25, ≥25 to <35, ≥35 kg/m^2)
- Admitting diagnosis (HF with reduced ejection fraction (LVEF ≤ 45%), acute respiratory insufficiency or acute exacerbation of COPD, acute ischemic stroke, acute infectious disease, and inflammatory diseases)
- History of diabetes (yes vs no)
- History of cancer (yes vs no)
- Baseline D-dimer level (> 2xULN vs ≤ 2xULN)
- Baseline aspirin use (yes vs no)
- Baseline thienopyridine use (yes vs no)
- Baseline proton pump inhibitor use (yes vs no)
- Duration of index hospitalization (3 to 6, 7 to 10, > 10 days)
- Duration of anticoagulation during index hospitalization (3 to 6, 7 to 10, > 10days)
- LMWH or UFH during index hospitalization (LMWH vs UFH)
- Use of US approved thromboprophylaxis (ie UFH, enoxaparin and dalteparin) vs other
- Total modified IMPROVE risk factor score (2, 3, ≥4)

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

An interim analysis for futility will be conducted when approximately 50% of subjects with adjudicated primary efficacy outcome have been observed (about 50% of the expected total number of events based on the actual observed blinded pooled event rate even if this event rate is lower than the assumed rate).

There will be no alpha adjustment to the final analysis. The IDMC will review results of the planned interim analysis and make a recommendation whether the study should be continued as planned, modified, or terminated prematurely due to futility or safety.

The study may be stopped early for futility when it would be unlikely to establish superiority on the primary efficacy outcome and/or a positive benefit over risk of rivaroxaban compared with placebo, if the study were to run to completion. The IDMC are encouraged to make a judgment for early stopping for futility when any one of 3 conditions below meets:

- Estimated HR on the primary efficacy outcome is 1.0 or higher based on the interim analysis. (Refer to Table 2 below for the corresponding conditional power)
- Observed primary efficacy outcome incidence rate by Day 45 in placebo at IA is too low that findings are not likely clinically meaningful
- Observed incidence rate of major bleeding on-treatment in rivaroxaban at IA is high enough that the overall benefit-risk profile would be unlikely to be positive at the final analysis as discussed in section 7

Table 2: Conditional Power (CP) Based on the Assumption of the Various Underlying RRRs in the Remaining Study.

	RRR in the Remaining Study	CP
Estimated HR=1 at IA 90% CI: (0.69, 1.44) 70% CI: (0.79, 1.26) 50% CI: (0.86, 1.16)	40%	32%
	30%	12%
	20%	4%
	10%	1%
	0%	0.3%

4. SUBJECT INFORMATION

4.1. Demographics and Baseline Characteristics

Descriptive statistics by treatment group will be provided for the following baseline demographics and disease characteristics, and not limited to the following baseline demographics and disease characteristics:

- Age
- Sex
- Race
- Geographic region
- CrCl
- Body Mass Index

- Admitting diagnosis (HF with reduced ejection fraction, acute respiratory insufficiency or acute exacerbation of COPD, acute ischemic stroke, acute infectious disease, and inflammatory diseases)
- History of diabetes
- History of cancer
- Baseline aspirin use
- Baseline thienopyridine use
- Baseline proton pump inhibitor use
- Duration of index hospitalization
- Duration of anticoagulation during index hospitalization
- LMWH or UFH during index hospitalization
- Use of US approved thromboprophylaxis (ie UFH, enoxaparin and dalteparin) vs other
- Total modified IMPROVE risk factor score
- Baseline D-dimer level (> 2xULN vs ≤ 2xULN)

4.2. Disposition Information

A subject will be considered a completer of the double-blind treatment if his/her last dose is at least at Day 45, at death date, or at primary efficacy outcome date. The number of non-completers of the double-blinded treatment and the reason for discontinuation will be summarized by treatment group. The distribution of time from the randomization to the last dose date will be shown by Kaplan-Meier plot.

The primary objective of this study will be addressed by primary efficacy outcome up to Day 45. Therefore, a subject will be considered a non-completer of the study if the subject's last outcome-evaluation date is before Day 45 date and the subject doesn't have a primary efficacy outcome or death before Day 45. The number of non-completers of the study and the reason for discontinuation will be summarized by treatment group. The distribution of time from the randomization to the last outcome-evaluation date will be shown by Kaplan-Meier plot.

A diagram similar to CONSORT Statement 2010 Flow Diagram[1] (Schulz et al, 2010) will be provided. A template of the diagram can be found in Attachment 1.

4.3. Treatment Compliance

For each subject, the treatment compliance rate is estimated as follows

Compliance rate (%) = 100 * actual treatment duration / intended treatment duration

More specifically, the actual treatment duration will be calculated by minimum of [Day 45 date, last dose date, and the first primary efficacy outcome date] - first dose date + 1 - dose interruption days up to Day 45. The intended treatment duration will be calculated by minimum of [Day 45 date, death date, and the first primary efficacy outcome date] - first dose date + 1 - dose interruption days up to Day 45 due to the reasons specified in protocol 10.2.1 (for example, hospitalization, bleeding, and other AEs).

Compliance rate will be summarized by treatment group.

4.4. Extent of Exposure

The number of subjects who take at least one dose of study drug will be provided for each treatment group.

Subject level duration of treatment exposure is defined as the duration between the first dose date and last dose date. Note that temporary study drug interruptions are included in this definition.

Duration of treatment exposure will be summarized by treatment.

4.5. Protocol Deviations

Protocol deviations will be summarized by treatment group. The categories of protocol deviations may include, but are not limited to, the following:

- Did not meet the inclusion or exclusion criteria
- Received wrong treatment or incorrect dose
- Did not discontinue study drug permanently according to the protocol
- Had been taking prohibited concomitant therapies as specified in the protocol

4.6. Prior and Concomitant Medications

Medication will be coded using the Anatomical Therapeutic Chemical (ATC) dictionary. The prevalence of medication use by ATC term will be summarized by treatment group and by study phase (prior to the first dose of study drug, during the study drug therapy period, and after last dose of study drug [post-treatment]). A medication that starts prior to the first dose of study drug is considered as a prior medication. A medication with a use period overlapping with the period between the first and last dose of study drug is considered as a concomitant medication. A medication that starts after the day of the last dose of study drug is considered as a post-treatment medication. Note that a medication use can be categorized into multiple study phases if its use spans into more than one study phase. In such a case, the medication use will be included in more than one summary.

5. EFFICACY

Unless otherwise stated, interim and final efficacy analyses will be based on CEC adjudicated events.

5.1. Analysis Specifications

5.1.1. Level of Significance

Unless otherwise specified, all statistical tests will be interpreted at a 2-sided nominal significance level of 0.05 and all confidence intervals at a 2-sided level of 95%.

5.1.2. Data Handling Rules

If the date of an outcome event is missing or partially missing, the earliest logically possible date after randomization will be used as the event date. The date of an outcome event is defined as the maximum of randomization and the first day of month if only day is missing and is defined as the first day of the following month after randomization if day and month are missing.

5.2. Primary Efficacy Outcome

5.2.1. Definition

The primary efficacy outcome is the composite of all symptomatic VTE events (lower extremity DVT, non-fatal PE) and VTE-related death (death due to PE or death in which PE cannot be ruled out as the cause).

5.2.2. Analysis Methods

The primary efficacy outcome will be analyzed based on the time from randomization to the first occurrence of symptomatic VTE and VTE-related death in the ITT analysis set and Up-to-Day-45 analysis phase. The primary statistical hypothesis will be tested using Cox proportional hazards model, stratified by subjects with CrCl from ≥30 to <50 mL/min versus subjects with CrCl ≥50 mL/min, with the treatment (as randomized) as the only covariate. Subjects will be analyzed according to the treatment group to which they are randomized, regardless of actual treatment received. As an operational arrangement, the 2-sided p-value will be reported, and if it is less than the 2-sided alpha of 0.05, then superiority of the study drug will be declared when the observed survival functions favor rivaroxaban. Note that the above testing procedure is equivalent to stratified logrank test[2]. The use of the Cox model format facilitates a consistent presentation of tests and CIs from analyses of efficacy and safety outcomes and from subgroup and sensitivity analyses.

The point estimate and corresponding 95% CI for the HR (rivaroxaban to placebo) will be provided based on the Cox proportional hazards model. For the CIs, the plausibility of proportional hazards assumption will be assessed by visually comparing the plot of the log of cumulative hazard between treatments, and additionally checked by adding a treatment by logarithm-transformed time interaction into the Cox model. If there is strong indication that the proportional hazards assumption is not valid, the 95% CI for relative risk of rivaroxaban versus placebo based on the incidence of the primary efficacy outcome from randomization up to Day 45 will be additionally provided, using the Cochran–Mantel–Haenszel (CMH) statistic with stratification by CrCl groups.

The cumulative event rate derived from the Kaplan-Meier method will be displayed graphically to evaluate the treatment effect over time.

The following sensitivity analyses of the primary efficacy outcome will be conducted to assess the robustness of the conclusion from primary efficacy analysis: a) stratified Cox model as in the primary efficacy analysis but additionally stratified by region, b) stratified Cox model as in the primary efficacy analysis but based on the per-protocol analysis set and on-treatment analysis phase. The additional sensitivity analyses of the primary efficacy outcome may be necessary to exclude subjects (e.g., significant quality issues, GCP violations, data integrity issues, regulatory agency commitment). Additional post-hoc analyses may be conducted to investigate unexpected

results. Off-treatment primary efficacy outcome events, defined by the events from the 3 days after the last dose date, will be summarized by treatment groups.

Extensive efforts will be made to collect complete event and vital status data for all subjects randomized in this study. The amount of unavoidable missing follow-up in this study with relatively short treatment duration of 45 days is not expected to influence the conclusion from the efficacy analyses. To illustrate the extent of missingness for the primary efficacy outcome, the numbers of subjects with missing follow-up and the percentages of total missing follow-up time out of total supposed follow-up time will be summarized. A non-completer of the study will be considered as a missing follow-up subject. Missing follow-up time will be calculated by Day 45 date - last outcome-evaluation date for a missing follow-up subject. For a subject, supposed follow-up time will be calculated by minimum of [Day 45 date, death date, first primary efficacy event date] - randomization date +1. Sensitivity analyses will be conducted to estimate any additional events from the rivaroxaban and placebo group during the missing follow-up time that are needed to overturn a superiority conclusion from the primary efficacy analysis. Similar sensitivity analyses may be conducted for each component of the primary efficacy outcome.

Homogeneity of treatment effects, both in HR and direction, in the subgroups defined in section 2.4 will be explored via a test for treatment by subgroup interaction in the stratified Cox model with treatment, subgroup and the treatment-by-subgroup interaction as the covariates, stratified by CrCl group (≥30 to <50, ≥50 mL/min) (at a 2-sided significance level of 0.05). For CrCl subgroups, unstratified Cox model will be used. Lack of a significant interaction will imply that the results are consistent across subgroups. If a significant interaction is observed, the results will be examined to determine whether the interaction is qualitative in nature using Pan and Wolfe approach[3] (1997). If the interaction is qualitative in nature, clinical explanations of the significant interaction will be explored.

For each individual subgroup, the estimate and CI for HR will be constructed based on the stratified Cox model except for the CrCl subgroups. If the number of subjects with a primary efficacy outcome in a treatment-by-stratum cell in a subgroup is less than 5, interpretation of the CI needs to be done with caution.

5.3. Secondary Efficacy Outcomes

5.3.1. Definition

The secondary efficacy outcomes are the following:

- VTE-related death (death due to PE or death in which PE cannot be ruled out as the cause)
- Symptomatic VTE (lower extremity DVT and non-fatal PE)
- The composite of symptomatic VTE (lower extremity DVT and non-fatal PE) and all-cause mortality (ACM)
- The composite of symptomatic VTE (lower extremity DVT and non-fatal PE), myocardial infarction (MI), non-hemorrhagic stroke and cardiovascular (CV) death (death due to a known CV cause and death in which a CV cause cannot be ruled out; by this definition, a VTE-related death is considered a CV death)
- All-cause mortality (ACM)

5.3.2. Analysis Methods

Each secondary outcome will be analyzed based on time from randomization to the first occurrence in the ITT analysis set, and the Up-to-Day-45 analysis phase. The same stratified Cox model as that for the primary efficacy outcome will be used.

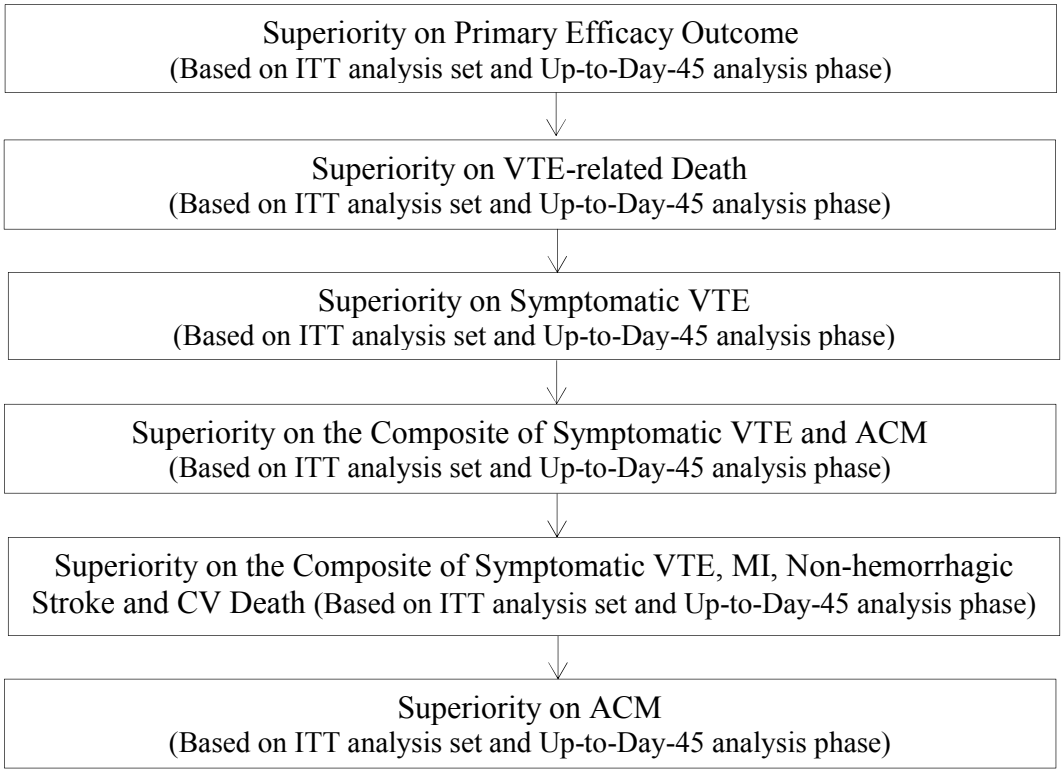
Other analyses described above for the primary efficacy outcome may be performed for secondary efficacy outcomes.

5.4. Multiple Testing Procedure

To control the family-wise type I error rate at alpha of 0.05 (2-sided) in testing for efficacy outcomes, if superiority of rivaroxaban over placebo on the primary efficacy outcome is established, superiority of rivaroxaban over placebo on secondary outcomes will be tested sequentially using a closed testing procedure in the hierarchical order as listed in 5.3.1, each at alpha of 0.05 (2-sided). If an individual test during any step is not statistically significant, later tests will not be declared to be statistically significant.

The following is a diagram of the multiple testing procedure.

Diagram of Multiple Testing Procedure



5.5. Other Efficacy Outcomes

5.5.1. Definition

Exploratory efficacy outcomes are:

- Symptomatic lower extremity DVT
- Symptomatic non-fatal PE
- Symptomatic upper extremity DVT
- MI
- Non-hemorrhagic stroke
- Re-hospitalization for symptomatic VTE (lower extremity DVT and non-fatal PE) within 30 days after randomization

Fatal MI or fatal stroke are defined as death occurred within 30 days of the MI or stroke and death was directly related to the MI or stroke respectively.

5.5.2. Analysis Methods

Each exploratory outcome will be summarized by treatment groups based on the ITT analysis set and Up-to-Day-45 analysis phase.

6. SAFETY

Unless otherwise stated, all safety summaries and analyses will be performed based on the safety analysis set. Subjects will be analyzed according to study drug received.

6.1. Bleeding Events

Unless otherwise stated, interim and final analyses of the principal safety outcome and non-major clinically relevant bleeding will be based on CEC adjudicated events.

The principal safety outcome is ISTH major bleeding. Treatments will be compared using the same Cox proportional hazards model as that for the primary efficacy outcome described earlier. The analysis will be based on the safety analysis set and on-treatment analysis phase. Treatments will also be compared using Anderson-Gill model for the time to multiple occurrences of the principal safety outcome. The analysis will be based on the safety analysis set and on-treatment analysis phase.

The homogeneity of treatment effects on the first occurrence of the principal safety outcome across subgroups defined in section 2.4 will be examined (at a 2-sided significance level of 0.05) via the test used for the primary efficacy outcome, based on the safety analysis set and on-treatment analysis phase.

Off-treatment major bleeding events, defined by the events from the 3 days after the last dose date, will be summarized by treatment groups.

Other safety outcomes are non-major clinically relevant bleeding and other bleeding.

Non-major clinically relevant bleeding will be analyzed based on time from randomization to the first occurrence. The same analysis as that for the principal safety outcome will be used.

All bleeding, other bleeding, bleeding by BARC classification as well as sub-categories of major bleeding, including critical sites, will be also summarized by treatment groups.

6.2. Adverse Events

For the purposes of this study (and after discussion with appropriate regulatory agencies) a value-driven approach to safety data collection will be utilized. Certain non-serious adverse events will not be collected, while certain events will be collected as outcome and therefore not reported as serious adverse events. All data on safety and outcomes will be reviewed regularly by an unblinded IDMC. For AEs that are collected, the verbatim terms reported in the eCRF by investigators to identify adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA).

Treatment-emergent adverse events (TEAE) are defined as those adverse events that start between the first study drug and the last study drug + 2 days.

Summary of the following adverse events will be provided by treatment group:

- Post-baseline (on or after the first study drug) adverse events
- Treatment-emergent adverse events
- Adverse events with onset > 2 days from the stop of study drug
- Post-baseline serious adverse events
- Treatment-emergent serious adverse events (TESAE)
- Serious adverse events with onset > 2 days from the stop of study drug
- Adverse events leading to permanent study drug discontinuation
- Adverse events with outcome of death

In addition, incidences of some of the above adverse events by system organ class and dictionary-derived (preferred) term will be provided.

Additional summaries, listings, or subject narratives may be provided, as appropriate.

6.3. Clinical Laboratory Tests

Because this study will not collect laboratory data routinely, laboratory data will not be summarized. Local lab data may be discussed in subject narratives.

6.4. Vital Signs and Physical Examination Findings

Descriptive statistics of pulse, blood pressure, and findings will be provided.

7. BENEFIT/RISK ASSESSMENT

This section describes several key elements for benefit-risk assessment, including 1) quantification method for benefits and risks, 2) efficacy and safety outcomes included in the evaluation, 3) analysis set and analysis phase, and 4) reporting format of the results.

These analyses are primarily intended to structure an integrated evaluation of the key benefits and risks in the study. They are complementary to other efficacy, safety analyses described elsewhere in this document. The overall benefit-risk profile of the study drug will be interpreted in consideration of the analyses described here and the totality of the data.

Quantification Methods for Measuring Benefits and Risks

As noted in previous sections, the treatment RRRs for efficacy and safety will be assessed using hazard ratios (Cox proportional hazards model). Because of difference in background event rate across different types of outcome events, a preferred metric used to evaluate treatment difference for benefit-risk assessment purpose is risk difference. For the current benefit-risk assessment, the treatment comparison of rivaroxaban vs. placebo will be evaluated based on the excess number of events between treatments, including events intended to be reduced (benefits) and events that may be caused (risks). The excess number of events is defined as the difference in event rates times a hypothetical population size (e.g., 10,000 patients). The event rate will be calculated using the following approaches:

- Event rate, expressed as the percentage of patients developing the event
- Kaplan-Meier Product-Limit estimates of cumulative event rates for the duration of study

In addition, number-needed-to-treat to benefit (NNT) or harm (NNH) also will be used to quantify benefits and risks of the treatment, respectively, which are calculated as the reciprocal of the differences in corresponding event rates.

Efficacy and Safety Outcomes for Benefit-Risk Evaluation

The efficacy and safety outcomes that will be included in benefit-risk evaluation are generally consistent with those specified in the efficacy and safety sections of this document, with some minor modifications as noted below.

A series of analyses will compare excess numbers of events for several pairs of key, mutually exclusive efficacy and safety endpoints:

- Composite of symptomatic VTE (lower extremity DVT, non-fatal PE) and VTE-related death (death due to PE or death in which PE cannot be ruled out as the cause) vs. ISTH major bleeding
- Composite of non-fatal PE and VTE-related death vs. composite of critical site bleeding and fatal bleeding
- VTE related death vs. fatal bleeding

To enable assessing the benefit-risk profile as a whole with mutually exclusive outcomes, we will also examine excess numbers of events for the following sets of outcomes:

- Benefits
    - The composite of VTE-related death, fatal MI and fatal ischemic stroke
    - Non-fatal ischemic stroke
    - Non-fatal MI
    - Non-fatal symptomatic PE
    - Symptomatic lower extremity DVT
- Harms
    - Fatal bleeding
    - Non-fatal critical site bleeding
    - Non-fatal, non-critical site bleeding requiring transfusion ≥2 units and non-fatal, non-critical site Hgb drop ≥ 2 g/dL
    - Non-fatal, non-critical site bleeding requiring transfusion ≥2 units
    - Non-fatal, non-critical site Hgb drop ≥ 2 g/dL
    - Non-fatal, non-critical site Hgb drop ≥ 2 g/dL not associated with transfusion ≥2 units

In addition, other major efficacy and safety outcomes identified in the protocol will also be evaluated in a similar fashion to support the overall benefit-risk evaluation.

Analysis Set and Analysis Phase

The primary analysis for benefit-risk evaluation will be based on the ITT analysis set and the Up-to-Day-45 analysis phase. However, additional analyses using analysis sets defined by other subject analysis sets and/or analysis phases as defined in the protocol or in this SAP may be performed as supportive analyses.

The benefit-risk analyses are not intended for hypothesis testing. Therefore, no multiplicity adjustment will be applied. When 95% CIs for point estimates of the excess number of events are provided as appropriate, nominal statistical significance at the alpha level of 0.05 (2-sided) will be declared if the confidence interval excludes zero.

The benefit-risk assessment will also be explored for the subgroups specified in Section 2.4.

A bi-variate analysis may be explored as post hoc analysis.

Reporting Format of the Results

To facilitate the comparison and interpretation of the results, data will be presented in one of the following formats as appropriate:

- Table format showing the between-treatment differences in benefits and risks (e.g., excess number of events and NNT or NNH)
- Kaplan-Meier plots depicting between-treatment differences in benefits and risks over time
- Forest plots for comparing key benefits and risks, as well as other key efficacy and safety outcome measures

8. PHARMACOKINETICS

Statistical summary of concentration (e.g. geometric mean, standard deviation, median, 5th and 95th percentile, range) by nominal sampling time window as specified in protocol will be provided. Statistical summary will also be stratified by CrCl groups.

The population PK analysis will be done to predict the overall exposure of rivaroxaban and estimate the PK model parameters for medically ill patients and to describe the effect of dosing modification on exposure for subjects with CrCl between ≥30 and <50 mL/min.

Results will be provided separately from the clinical study report.

Separate analysis plan document will be provided.

9. HEALTH ECONOMICS

For medical resource utilization and health economics, exploratory outcomes include re-hospitalization (including LOS), ER visits, ICU and CCU stays, including LOS and subject discharge destination (after re-hospitalization).

Each exploratory outcome will be descriptively summarized by treatment group.

REFERENCES

1. Schulz KF, Altman DG, Moher D, for the CONSORT Group (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ: 340:c332 (http://www.bmj.com/content/340/bmj.c332).
2. Therneau, Terry M. and Patricia M. Grambsch (2000). Modeling Survival Data: Extending the Cox Model. Statistics for Biology and Health. Springer, New York.
3. Guohua Pan and Douglas A. Wolfe (1997). Test for Qualitative Interaction of Clinical Significance. Statistics in Medicine, Vol. 16, 1645-1652.

ATTACHMENTS

Attachment 1: Template of CONSORT Statement 2010 flow diagram