Janssen Research & Development

Statistical Analysis Plan

A Randomized, Double-blind, Event-driven, Multicenter Study Comparing the Efficacy and Safety of Rivaroxaban with Placebo for Reducing the Risk of Death, Myocardial Infarction or Stroke in Subjects with Heart Failure and Significant Coronary Artery Disease Following an Episode of Decompensated Heart Failure

Protocol RIVAROXHFA3001; Phase 3

JNJ-39039039; BAY 59-7939

Amendment 2.0

Status:	Approved
Date:	12 December 2017
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-ERI-54766356, 4.0

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

TABLE OF CONTENTS

SAP AMENDMENT

SAP Version	Issue Date
Original SAP	June 10, 2013
Amendment 1.0	February 01, 2017
Amendment 2.0	December 12, 2017

Amendments are listed beginning with the most recent amendment.

Amendment 2.0

The principal rationale for the SAP Amendment 2.0 is the addition of a supporting analysis suggested by the trial's Steering Committee, and to clarify the definition of an analysis subpopulation. Secondarily, the Amendment will also add a recently approved medication to the list of baseline therapies, address differences in adverse event (AE) reporting in one country (Japan), clarify text and correct minor grammatical errors.

Applicable Section(s)	Description of Change(s)
Rationale: To add a supporting analysis suggested by the trial's Steering Committee**.**	
5.3.2. Analysis Methods	Added the sentence "A supporting analysis of time to the first occurrence of all-cause mortality (ACM) or re-hospitalization for worsening of HF will also be performed using the same analysis method as for the primary efficacy endpoint.".
Rationale: To clarify the definition of a subpopulation**.**	
2.4. Definition of Subgroups; 5.5. Additional Efficacy; Analyses; 6.1. Bleeding Events	Changed the phrase "subjects who complete treatment at GTED and are scheduled for EOS visit (Office/Phone)" to "Subjects who complete the double-blind treatment period and have the EOS visit (Office/Phone)"
Rationale: To add a new medication "angiotensin receptor-neprilysin inhibitor" (ARNI) to the list of baseline therapies. ARNIs are currently being coded to ANGIOTENSIN II ANTAGONISTS, OTHER COMBINATIONS in this study since this term is not yet available in the current WHODRUG coding dictionary.	
ABBREVIATION; 4.1. Demographics and Baseline Characteristics	Added "ARNI, angiotensin receptor-neprilysin inhibitors" to the abbreviation list. Added "Baseline ARNI use" as a sub-bullet of "Baseline ARB use" in the list of demographics and baseline characteristics to be summarized.
Rationale: To reflect the current Standardized Medication Name and Medication Class in the WHODRUG coding dictionary for spironolactone and epleronone.	
2.4. Definition of Subgroups; 4.1. Demographics and Baseline Characteristics	Replaced "mineralocorticoid receptor antagonist" with "aldosterone antagonists".
Rationale: To clarify the definition of the ITT population i.e., the same subject randomized twice will only be counted once in the ITT population.	
2.3. Analysis Sets	Added the word "unique" into the sentence of: Intention-to-treat (ITT) population: This subject population consists of all randomized unique subjects who have a signed valid informed consent.

Applicable Section(s)	Description of Change(s)
Rationale: To add fragile subjects as a subgroup.	
2.4. Definition of Subgroups	Added a subgroup: • Fragile subjects (yes vs no; fragile subjects are those with age > 75 years or weight ≤ 50 kg or baseline eGFR < 50 mL/min)
Rationale: No more countries will be added to the study.	
2.4. Definition of Subgroups	Deleted the text "(Regions will be updated if additional countries are included in this study.)".
Rationale: To clarify the analysis sets used for evaluation of the primary and secondary efficacy endpoints.	
2.3.1.1. Main Efficacy Analysis Set; 2.3.1.2, Sensitivity Efficacy Analysis Set	Changed the word "Primary" in the section heading, "Primary Efficacy Analysis Set", to "Main" and the word "Secondary" in the section heading, "Secondary Efficacy Analysis Sets", to "Sensitivity".
Rationale: To describe the additional analysis sets previously defined in multiple places in the document in one section.	
2.3.1.3. Additional Analysis sets	Added a separate Section (2.3.1.3) to describe the additional analysis sets.
Rationale: To clarify the definitions of the observation periods used for the safety analyses.	
2.3. Analysis Set; 2.3.2. Safety Analysis Set	Added the paragraphs: "The main safety analysis set is defined by the On-treatment observation period and the safety subject population. The sensitivity safety analysis set is defined by the post-first dose of study drug observation period and the safety subject population. Subjects will be analyzed according to the treatment assigned."; "Bleeding events and adverse events will be analyzed and summarized based on both the main and sensitivity safety analysis sets."; "In addition, the same additional analysis sets defined in Section 2.3.1.3 will also be used for the summaries of safety events. Details are described in Section 6.1.". Moved the paragraph below from Section 2.3.2 to Section 2.3: "Post-first dose of study drug: This observation period starts from the day of the first dose of study drug and ends on the day of last contact, inclusively. Because subjects will be asked to take their first dose of study drug on the randomization day, while they are still in investigators' offices, for most of subjects, this observation period will be identical to post-randomization observation period.". Deleted the two paragraphs: "In addition to observation periods defined in Section 2.3.1, the following observation period is defined specifically for the safety analysis."; "Bleeding events will be analyzed based on the analysis set defined by the on-treatment observation period and the safety subject population. Subjects will be analyzed according to the treatment assigned. As a sensitivity analysis, bleeding events will also be analyzed based on the analysis set defined by the post-first dose of study drug observation period and the safety subject population. AEs reported by investigators will be summarized based on the analysis set defined by the on-treatment observation period and separately by the post-first dose of study drug observation period and the safety subject population.".

Applicable Section(s)	Description of Change(s)
Rationale: To clarify the definition of a baseline therapy.	
4.1. Demographics and Baseline Characteristics	Added the text "or implantable defibrillator or combination" to the bullet of "Cardiac resynchronization therapy".
Rationale: While the number of subjects who permanently discontinue study drug will be reported, the data collected are insufficient to support an analysis of the conversion from a temporary to a permanent study drug discontinuation**.**	
4.2. Disposition Information	Deleted the two paragraphs: "It is possible that a temporary study drug interruption can become a permanent study drug discontinuation due to events that occur after the start of the temporary study drug interruption, such as diagnosis of atrial fibrillation that requires permanent anticoagulant therapy or announcement of GTED while a subject is on temporary study drug stoppage. The number of subjects who permanently discontinue the study drug after a temporary study drug interruption may be summarized if there are sufficient observations. "The same method will be used to summarize information of discontinuations from the study if data is sufficient."
Rationale: To clarify the wording**.**	
2.3.1. Analysis Sets	Clarified the definition of the "Per-Protocol (PP) population: This population is a subset of the ITT population, consisting of subjects who do not have a major protocol deviation in the following categories …".
5.1.1. Level of Significance	Clarified the main, sensitivity and additional analysis sets for the efficacy, principal safety (bleeding) and adverse event analyses Replaced the sentence "If p value for a hypothesis is greater than that significant level, all remaining hypotheses after this hypothesis in the fixed sequence will not be rejected." with "If an individual test during any step is not statistically significant, further treatment comparison may continue (i.e., reporting of p-values) but significance will not be claimed.".
Rationale: To reflect the planned missing data analysis.	
5.2.2. Analysis Methods	Added "on the primary efficacy analysis" to the first sentence of the paragraph starting "The potential impact of missing data …". Deleted the phrase "If the missing observation are substantial," in the same paragraph.
Rationale: To describe additional analyses for Japan subjects based on Japan-specific reporting of bleeding events and adverse events.	
6.1. Bleeding Events	Added a sentence: "For Japan subjects only, summary statistics will be provided for non-major clinically relevant bleeding events and minimal bleeding events.".
6.2. Other Adverse Events	Added a paragraph "Based on specific safety reporting requirements from the Pharmaceuticals and Medical Devices Agency, AE data from Japan may not be directly comparable with those collected from non-Japan sites. Therefore, the AE data from Japan sites may be summarized separately from AE data from non-Japan sites in the respective reported safety tables.".
Rationale: To reflect the intent of the protocol and current data reporting requirements.	

Applicable Section(s)	Description of Change(s)
6.4. Vital Signs and Physical Examination Findings	Revised the section wording to "Targeted medical history and vital signs are obtained at baseline only and summarized as baseline characteristics. Physical examinations are not required as part of study procedures. Available data may be discussed in subject narratives.".
Rationale: To reflect that a fatal or non-fatal MI and a fatal or non-fatal ischemic stroke may be counted in the Benefit outcome.	
7. Benefit/Risk Assessment	Added a sentence to the Benefit outcome discussion "As a supporting measure, the composite of non-bleeding related death, MI (fatal, non-fatal) or ischemic stroke (fatal, non-fatal) may also be evaluated.".
Rationale: To move the text to the appropriate subsection due to formatting error in previous version**.**	
5.1.1. Leve of Significance; 5.1.2. Data Handling Rules	Moved the text "Confidence intervals (CIs) are calculated at the 95% nominal confidence level. In addition, repeated CI (RCI), which adjusts for the interim analysis, will also be provided in the clinical study report (CSR)" to the bottom of Section 5.1.1.
2.2. Pooling Algorithm for Analysis Centers; 2.3 Analysis Set	Moved the text "No pooling algorithm for analysis centers will be used. The primary efficacy analysis will be stratified by region defined in Section 2.4." to Section 2.2.
Rationale: Consistency and minor errors were noted.	
Throughout the SAP	Consistency with protocol and minor grammatical, spelling or formatting changes were made.

Amendment 1.0

The overall rationale for the SAP Amendment 1.0: is due to changes in the Study Protocol (Amendments 1, 2, 3) and FDA feedback on the planned statistical methods.

Applicable Section(s)	Description of Change(s)
Rationale: To indicate that changes made in this document are based on changes made in the protocol amendments**.**	
1. Introduction	Added the texts "originally" and "all its amendments" in the sentence "This SAP is based on …".
Rationale: To include high-risk patients treated in an outpatient setting with parenteral medications for decompensated heart failure.	
1.1. Trial Objectives	Replaced the phrase "following a recent hospitalization for exacerbation of HF" with "following an episode of decompensated HF (index event)". Replaced the phrase "a recent hospitalization for exacerbation of HF" with "index event".
Rationale: To clarify the definition of the bleeding events requiring hospitalization.	
1.1. Trial Objectives	Added the text "(caused a hospitalization or prolongation of an existing hospitalization)".
Throughout the document where the study population is described	Deleted the term "chronic" or replaced "chronic" with "symptomatic". Replaced the phrase "a recent hospitalization for exacerbation of HF" with "an episode of decompensated HF".
Rationale: Protocol Amendment INT-3 increases the target number of primary efficacy events from 984 to 1,200 and increases the target number of primary efficacy events from 500 to 600 for the interim analysis (IA).	

Applicable Section(s)	Description of Change(s)
1.2 Trial Design; 2.1. Subject-Level Trial Milestone Dates	Replaced "984" with "1,200".
1.2 Trial Design; 1.4. Sample Size Justification; 3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW; 5.2.2. Analysis Methods	Changed "500±50" to "approximately 600".
1.4. Sample Size Justification	Added: The study is revised to observe occurrences of the primary efficacy event in 1,200 unique randomized subjects based on a simulation study performed on currently available trial data, i.e., longer enrollment period, lower than anticipated primary efficacy event rate, higher than anticipated study drug discontinuation rate: • Event rate in the placebo arm: approximately 16%/year based on the observed combined event rate at 14%/year (assumed 20% of RRR remains) • Early permanent study drug discontinuation rate: approximately 13%/year • Duration of enrollment period: approximately 48 months (projected) The simulation study estimated that the duration of study may extend to approximately 54 months based on the fact that the original goal of enrolling a total of 5,000 subjects remains unchanged and all subjects would be observed for approximately 6 months after the last subject is enrolled. With a 20% effect size and overall α level of 5% (2-sided), 1,200 number of primary efficacy events will produce a statistical power of approximately 80%. Deleted the sentence "The interim analysis will have approximately 75% power to detect an RRR of 30% in the primary efficacy endpoint."

Rationale: To be consistent with the wording used in Protocol Amendment INT-3.

Applicable Section(s)	Description of Change(s)
1.2. Trial Design	Deleted the sentences "The screening phase will last up to 28 days (up to 21 days during the index hospitalization and up to 7 days after discharge from the index hospitalization), followed by an estimated 6 to 30 months double-blind treatment phase." Added the sentences "The screening phase may last up to 51 days before randomization (21 days maximum for a hospitalization admission plus up to 30 days after discharge from the facility treating the index event), followed by an estimated 6 to 54 months of double-blind treatment phase." and "Note that the screening period in Japan may last up to 60 days before randomization (up to 30 days for a hospitalization admission plus up to 30 days after discharge from the facility treating the index event)." Deleted the sentence "If the rate of efficacy events is lower than the estimated used in sample size calculation, the sample size may increase up to 5500 subjects." Added a phrase "as soon as possible after the last dose of study drug" at the end of sentence "Subjects who permanently discontinue study drug …". Deleted the sentence "A subject will be instructed to discontinue study drug temporarily for an outcome event such as an MI or ischemic stroke." Added the sentences: "If a subject experiences an outcome event such as MI or stroke, study drug may be temporarily held if necessary or if using other anticoagulant drugs or thrombolytic therapy" and "This also applies to other efficacy outcome events such as DVT or PE." Added the phrase "or for any reason listed in Section 10.2.2 of the protocol." at the end of sentence "After appropriate treatment of the event, …".
Rationale: The process of outcome event document verification has evolved over time.	
1.2. Trial Design	Deleted the sentences "Unless otherwise specified, outcome events hereafter refer to events with adequate supporting documentation per protocol; and all-reported events refer to outcome events reported by investigators, regardless of adequacy of supporting documentation." and "Supporting documentation will be reviewed, and the adequacy of documentation will be determined before database lock and captured in database." Added the sentence "The details of outcome event reporting and verification are described in the Investigator Manual for Outcome Events and the Sponsor Outcome Event Verification Process Manual."
Rationale: To clarify that the D-dimer collected at both baseline and Week 4 are not being shared with investigators nor with sponsor's clinical study team to avoid any bias in treating the subjects.	
1.5. Randomization and Blinding; 6.3. Clinical Laboratory Tests	Add the sentence "Complete D-dimer data will not be available until after the database lock." Deleted the paragraph "While the personnel at the study site may be aware of subjects randomly selected, Week 4 D-dimer value will only be available to the sponsor after data unblinding because of the potential of treatment assignment unblinding. Week 4 D-dimer results will not be given to the subject or the investigator. Therefore, collection of Week 4 D-dimer data will not unblind treatment assignment."
Rationale: To capture all relevant data collection for the last contact date.	

Applicable Section(s)	Description of Change(s)
2.1. Subject-Level Trial Milestone Dates	Added the sentence "If a subject completes the randomization visit but no further visits and no other data is recorded on the eCRF, the randomization date will be considered as the last contact date." Added the phrase "and the death date is missing" to the sentence "For subjects who died during the study ….".
Rationale: In line with MPD category text for subjects excluded from per-protocol analysis set.	
2.3.1. Efficacy Analysis Set(s)	Changed the phrase "Subjects with the following major protocol deviations will be excluded from the PP population:" to "Subjects with a subset of major protocol deviations in the following categories will be excluded from the PP population (see FRM-15949 for the specific major protocol deviations that are exclusions from the PP population and identification/review process):" Deleted: • Not meeting key inclusion or exclusion criteria • Taking incorrect study drug • Not permanently discontinuing study drug according to the protocol • Treatment compliance rate < 75% (see Section 4.3 for details) • Having been taking prohibited concomitant therapies as specified in the protocol for an extended period (≥ 30 consecutive days or 20% of treatment period on a continuous basis, whichever is shorter). A list of prohibited concomitant medications is in Attachment 3. Added: • Entered but did not satisfy criteria • Received a disallowed concomitant treatment • Received wrong treatment or incorrect dose (includes subjects with treatment compliance rate < 75% calculated according to a formula described in Section 4.3) • Developed withdrawal criteria but not withdrawn • Other (as appropriate)
Rationale: To change the safety analyses from "study drug received" to "study drug assigned" based on the fact that the COMMANDER HF is a long term study and the incidence of a subject receiving the wrong study drug during the entire treatment period will be low.	
2.3.2. Safety Analysis Set; 6.1. Bleeding Events	Change the sentence "Subjects will be analyzed according to the actual treatment received" to "Subjects will be analyzed according to the treatment assigned". Deleted the sentence "If a subject receives both rivaroxaban and placebo, the subject will be analyzed according to the treatment assigned".
Rationale: To be consistent with the subgroups defined in the protocol and to define regions with more countries.	
2. 4. Definition of Subgroups	Deleted "History of percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG) (yes vs no)", "Cardiac resynchronization therapy (yes vs no)", "aldosterone inhibitors", "White, Asian, Black, and others". Added "Baseline BNP or NT-proBNP (≤ median vs > median), "mineralocorticoid receptor antagonist". Expended regions with more countries enrolling subjects.
Rationale: To change the stratification factor for the efficacy and safety analyses from "country" to "region" due to the increase in the number of enrolling countries and the corresponding smaller sample sizes expected in those countries.	

Applicable Section(s)	Description of Change(s)
Section 3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW; 5.2.2 Analysis Methods; 6.1. Bleeding Events	Replace “stratified by pooled countries” with “stratified by region”.
Rationale: To add common baseline demographics and baseline characters appropriate for the study population.	
4.1. Demographics and Baseline Characteristics	Deleted “History of CABG”, “aldosterone inhibitors”. Added “Height”, “Weight”, “mineralocorticoid receptor antagonist use”, “Baseline BNP or NT-proBNP”. Added the phrase “Additionally, evidence of coronary artery disease will be summarized.”
Rationale: To plan for proper analyses for subject disposition information.	
4.2. Disposition Information	Added the sentence “The same method will be used to summarize discontinuation from the study.” Deleted the phrases: “These permanent early study drug discontinuations will be included in the summary of early permanent discontinuation of study drug with all other early permanent drug discontinuations. For these early permanent study drug discontinuations, an additional summary will be provided to show reasons of temporary study drug discontinuations immediately preceding these early permanent study drug discontinuations. Details are specified in the DPS.” Added the sentence “The number of subjects who permanently discontinue the study drug after a temporary study drug interruption may be summarized if there are sufficient observations.”
Rationale: Study drug compliance will be calculated based on the days of study drug taken due to difficulties of obtaining reliable pill counts information.	
4.3. Treatment Compliance	Replaced the compliance formula based on the pill counts information with the formula based on the days of study drug taken. Deleted the phrases: “If a pill count is impossible for a bottle due to reasons such as lost bottle or bottle not returned, it will be assumed that the compliance rate during the period that the bottle is intended to cover is the same as the compliance rate during the rest of the study. This is equivalent to excluding the period with missing pill count from the compliance rate calculation”. Added the sentence “If the dose date prior to a treatment interruption is missing, the same imputation rule for imputing the last dose date descripted in Section 2.1 will be used to impute the dose date.”
Rationale: The MPD category text was revised to reflect the database terminology.	

Applicable Section(s)	Description of Change(s)
4.5. Protocol Deviations.	Deleted: • Not meeting inclusion or exclusion criteria • Receiving incorrect medication kits • Not permanently discontinuing study drug as per protocol • Having been taking prohibited concomitant therapies for an extended period (≥ 30 consecutive days or 20% of treatment period on a continuous basis whichever is shorter). A list of prohibited concomitant medications is provided in Attachment 3. Added: • Entered but did not satisfy criteria • Received a disallowed concomitant treatment • Developed withdrawal criteria but not withdrawn (includes subjects with treatment compliance rate < 75% calculated according to a formula described in Section 4.3) • Received wrong treatment or incorrect dose • Other (as appropriate)
Rationale: To clearly indicate the stratification factor and analysis sets for sensitivity analyses.	
5.2.2 Analysis Methods	In the sentence "These include, but are not limited to, …", added the phrases "a log-rank test stratified by country and"; "for the analysis set defined by the up-to-GTED observation period and ITT subject population"; and "for the analysis set defined on-treatment observation period and the safety subject population". Added the sentence "The primary efficacy endpoint will also be analyzed for the ITT subjects who enrolled under the original protocol and who enrolled under the amended protocols for the Up-to-GTED observation period."
Rationale: To indicate that substantial missing observations will warrant sensitivity analyses.	
5.2.2 Analysis Methods	Reworded the sentence "Additional details of these sensitivity analyses, including specifications for imputation, will be provided in a separate document that will be finalized prior to planned interim analysis" to "If the missing observations are substantial, additional details of these sensitivity analyses, including specifications for imputation, will be provided in a separate document that will be finalized prior to the final database lock."
Rationale: To detail the definition of fatal bleeding and the analysis for other reported bleedings.	
6.1. Bleeding Events	Added the protocol fatal bleeding definition: "A fatal bleeding event is one in which the subject dies within 7 days of a bleeding event requiring hospitalization or ISTH major bleeding." Added 3 exclusive categories of fatal bleeding with regard to how a fatal bleeding event is captured on the bleed page of eCRF. Added the sentence "The number of the subjects with other reported bleeding events will be summarized by treatment group and bleeding site."
Rationale: To clarify the selected lab test collections during the study.	
6.3. Clinical Laboratory Tests	Added the phrase "during the study other than D-dimer at week 4 (10% of subjects) and hemoglobin at week 12 (up to week 24)" to the first sentence.
Rationale: The duration of the study is longer than anticipated	
7. Benefit/Risk Assessment	Added the sentence "*Because the duration of the trial is longer than anticipated, Kaplan-Meier Product-Limit estimates of cumulative event rates at 2-years and 3-years may be explored."

Applicable Section(s)	Description of Change(s)
Rationale: To address FDA's comments on the original study SAP (submitted to the Agency on June 14, 2013; FDA comments received on July 30, 2013).	
2.2. Pooling Algorithm for Analysis Centers	The original pooling strategy was deleted, i.e., countries would not be pooled. Added the sentence "No pooling algorithm for analysis centers will be used. The primary efficacy analysis will be stratified by region defined in Section 2.4."
2.3.1. Efficacy Analysis Set(s)	Added: Off-treatment: this observation period includes data from 3 days after the last dose of the study drug to the last contact date, inclusively. Day 3 to day 30 after the last dose: the 3rd day after the last dose of study drug to the 30th day after the last dose of study drug or the last contact day, whichever is earlier.
2.4. Definition of Subgroups	Two additional subpopulations were added for summarizing efficacy and safety endpoints: In addition, the number of subjects with efficacy endpoints including components of composite efficacy endpoints; the number of subjects with principal safety endpoint and its components of fatal bleeding (defined in Section 6.1) and bleeding into a critical space with a potential for permanent disability; and the number of subjects with other bleeding outcomes of ISTH major bleeding events and bleeding requiring hospitalization will be summarized for safety subjects in the following subpopulations: • Subjects who permanently discontinue study drug for any reason prior to GTED • Subjects who complete treatment at GTED and are scheduled for EOS visit (Office/Phone)
3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW	Changed the phrase "at 5%, 2-sided α level" to "at a significant level corresponding to the Pocock-type boundary" in the sentence "If the test statistic for the primary efficacy endpoint …, secondary endpoints will …".
5.1.1. Level of Significance; 5.3.2 Analysis methods	Changed the phrase "greater than 0.05" to "greater than that significant level" in the sentence "If p-value for a hypothesis is …". Changed the phrase "5% significance level" to "a significance level corresponding to the Pocock-type boundary". Changed '0.05 to "corresponding to the Pocock-type boundary". Deleted the sentence "In summary, all statistical tests will be conducted at the 0.05 significance level, except the test for the primary efficacy endpoint." Added the sentence "In addition, repeated CI (RCI), which adjusts for the interim analysis, will also be provided in the clinical study report (CSR)."
5.2.2. Analysis Methods	Added the sentence "In addition, a RCI will be calculated based on the procedure described in Jennison and Turnbull (2000), where significance level at the interim analysis and the final analysis will be approximately 99.7% and 95.2%, respectively."

Applicable Section(s)	Description of Change(s)
5.5. Additional Efficacy Analyses	Added the new section: 5.5. Additional Efficacy Analyses For all primary, secondary and other efficacy outcome events discussed above, the number of subjects with these events will be tabulated for the following analysis sets: • Off-treatment observation period for subjects who permanently discontinue study drug for any reason prior to GTED • Off-treatment observation period for subjects who complete treatment at GTED and are scheduled for EOS visit (Office/Phone) • Off-treatment observation period for the Safety population. • Day 3 to Day 30 after the last dose observation period for subjects who permanently discontinue study drug for any reason prior to GTED • Day 3 to Day 30 after the last dose observation period for subjects who complete treatment at GTED and are scheduled for EOS visit (Office/Phone) • Day 3 to Day 30 after the last dose observation period for the Safety population.
6.1. Bleeding Events	Added: The number of subjects with above safety outcomes will be tabulated for the following analysis sets: • Off-treatment observation period for subjects who permanently discontinue study drug for any reason prior to GTED • Off-treatment observation period for subjects who complete treatment at GTED and are scheduled for EOS visit (Office/Phone) • Off-treatment observation period for the Safety population. • Day 3 to Day 30 after the last dose observation period for subjects who permanently discontinue study drug for any reason prior to GTED • Day 3 to Day 30 after the last dose observation period for subjects who complete treatment at GTED and are scheduled for EOS visit (Office/Phone) • Day 3 to Day 30 after the last dose observation period for the Safety population.
Rationale: Consistency and minor errors were noted.	
Throughout the SAP	Consistency with protocol and minor grammatical, spelling or formatting changes were made.

ABBREVIATIONS

ACEI	angiotensin-converting enzyme inhibitor
ACM	all-cause mortality
AE	adverse event
ARB	angiotensin receptor blocker
ARNI	angiotensin receptor-neprilysin inhibitors
BNP	brain natriuretic peptide
CABG	coronary artery bypass graft
CAD	coronary artery disease
CDF	cumulative distribution function
CEC	Clinical Endpoint Committee
CI	confidence interval
CSR	Clinical Study Report
CV	Cardiovascular
DPS	Data Presentation Specifications
DVT	deep vein thrombosis
ECG	Electrocardiogram
EOS	End-of-Study
eCRF	electronic case report form
eGFR	estimated glomerular filtration rate
FDA	Food and Drug Administration
GCP	Good Clinical Practice
GTED	Global Treatment End Date
HF	heart failure
IDMC	Independent Data Monitoring Committee
ISTH	International Society on Thrombosis and Haemostasis
ITT	Intent-to-Treat
IWRS	interactive web response system
LVEF	left ventricular ejection fraction
MedDRA	Medical Dictionary for Regulatory Activities
MRU	medical resource utilization
MI	myocardial infarction
NT-proBNP	N-terminal-pro- brain natriuretic peptide
OE	Outcome Event
PCI	percutaneous coronary intervention
PE	pulmonary embolism
PP	Per-Protocol
RCI	Repeated confidence interval
RRR	relative risk reduction
SAP	Statistical Analysis Plan
SC	Steering Committee

1. INTRODUCTION

This statistical analysis plan (SAP) specifies definitions of analysis sets, key derived variables and statistical methods for analysis of efficacy and safety for the Phase 3 study RIVAROXAHF3001 (also known as COMMANDER HF). This SAP is based on the Clinical Protocol RIVAROXAHF3001, originally finalized on April 18, 2013 and all its amendments.

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 efficacy and safety results, and make a recommendation whether the study will be continued as planned, or terminated early for overwhelming superiority, futility, or safety. Safety data will be reviewed by the IDMC periodically and additional details for the planned interim analysis are specified in the abbreviated IDMC SAP, which was specifically prepared for IDMC use.

1.1. Trial Objectives

The primary objective is to demonstrate that rivaroxaban is superior to placebo in subjects with symptomatic heart failure (HF, 3 months or longer) and significant coronary artery disease (CAD), who are receiving standard care, in reducing the risk of the composite of all-cause mortality (ACM), myocardial infarction (MI), or stroke following an episode of decompensated heart failure (index event).

The secondary objectives are to compare rivaroxaban with placebo in addition to standard care in subjects with symptomatic HF and significant CAD following an episode of decompensated heart failure (index event) in reducing the risk of the following outcomes:

- Composite of cardiovascular (CV) mortality and re-hospitalization for worsening of HF
- CV mortality
- Re-hospitalization for worsening of HF
- Re-hospitalization for CV events

The exploratory objectives are to compare rivaroxaban with placebo in addition to standard care in subjects with symptomatic HF and significant CAD following an index event in reducing the risk of the following outcomes:

- Selected medical resource utilization (MRU) data on re-hospitalization for CV events and for worsening of HF
- Symptomatic deep vein thrombosis (DVT)
- Symptomatic pulmonary embolism (PE)
- Benefit-risk balance

Note that the plan for analyses of MRU data on CV and HF re-hospitalization will be provided in a separate document. They are not covered in this document.

The safety objectives are to compare the following bleeding events between rivaroxaban and placebo in addition to standard care in subjects with symptomatic HF and significant CAD following an index event:

- The composite of fatal bleeding or bleeding into a critical space (intracranial, intraspinal, intraocular [vitreous or retinal], pericardial, intra-articular, retroperitoneal, intramuscular (with compartment syndrome) with a potential for permanent disability
- Bleeding events requiring hospitalization (caused a hospitalization or prolongation of an existing hospitalization)
- Major bleeding events using the International Society on Thrombosis and Haemostasis (ISTH) bleeding criteria

Overall safety will also be assessed.

1.2. Trial Design

This is a randomized, double-blind, placebo-controlled, parallel-group, multicenter, event-driven, group-sequential, superiority study of rivaroxaban with clinical outcome assessments in subjects with symptomatic HF (3 months or longer) and significant CAD. The study consists of a screening phase, a double-blind treatment phase, and a follow-up after the sponsor-announced global treatment end date (GTED, defined as the date when 1,200 primary efficacy outcome events are predicted to have occurred (date is based on site local time). The screening phase may last up to 51 days before randomization (21 days maximum for a hospitalization admission plus up to 30 days after discharge from the facility treating the index event), followed by an estimated 6 to 54 months of double-blind treatment phase. Note that the screening period in Japan may last up to 60 days before randomization (up to 30 days for a hospitalization admission plus up to 30 days after discharge from the facility treating the index event). The follow-up after the GTED is the End-of-Study (EOS) visit, which occurs 30±15 days after GTED. The primary efficacy outcome is the composite of ACM, MI, or stroke. Outcome events will *not* be centrally adjudicated in this study. The decision of no central adjudication for outcome events is based on 2 considerations: 1) the majority of events are expected to be ACM, for which no adjudication is needed; and 2) the confirmation rate by the Clinical Endpoint Committee (CEC) for MI and stroke was high in the previous ROCKET AF and ATLAS ACS 2 TIMI 51 studies (ROCKET AF: 73% for MI and 85% for stroke; ATLAS ACS 2 TIMI 51: 85% for MI and 85% for stroke). Since outcome events are not centrally adjudicated, only outcome events with adequate supporting documentation per protocol will be included in the main analyses. Reported outcome events with inadequate documentation will be included in sensitivity analyses. The details of outcome event reporting and verification are described in Investigator Manual for Outcome Events and Sponsor Outcome Event Verification Process Manual.

Approximately 5,000 subjects who meet all inclusion criteria and none of exclusion criteria will be randomly assigned to the rivaroxaban 2.5 mg b.i.d. arm or the placebo b.i.d. arm in a 1:1 ratio. Randomization will be stratified by country. Randomized subjects are expected to remain in the

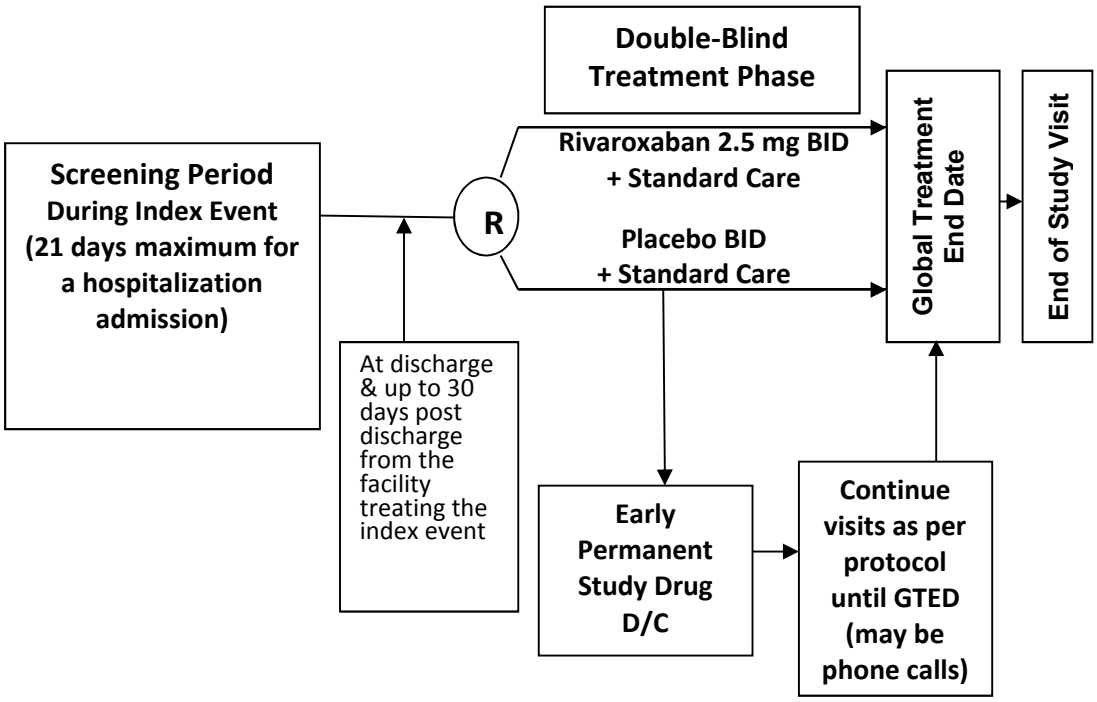
double-blind treatment phase until the sponsor-announced GTED, at which time subjects will be instructed to discontinue study drug and return to the study site 30±15 days after GTED for the EOS visit. At the EOS visit efficacy and safety outcomes will be collected.

Subjects who permanently discontinue study drug before GTED will complete the Early Permanent Study Drug Discontinuation visit as soon as possible after the last dose of study drug. In addition, these subjects should return for all scheduled visits, including the EOS visit. If these subjects refuse office visits, the investigator will be asked to encourage the subjects to allow regular contact (e.g., by telephone) until study end according to the Time and Events Schedule in the protocol, either with them, or with a close friend, relative, or their primary care physician to determine vital status and if an efficacy or safety outcome event has occurred.

A subject will be considered as having completed the double-blind treatment phase if the subject continues taking double-blind study drug until either the GTED or within 7 days before the death of the subject. If the last dose day cannot be determined in a subject who dies, the last dose day will be imputed according to the algorithm presented in Section 2.1, and the above definition will be applied to the imputed last dose day. If a subject experiences an outcome event such as MI or stroke, study drug may be temporarily held if necessary or if using other anticoagulant drugs or thrombolytic therapy. This also applies to other efficacy outcome events such as DVT or PE. After appropriate treatment of the event, the investigator may choose to resume study drug for the subject. Study drug must be permanently discontinued for any intracranial hemorrhage or for any reason listed in Section 10.2.2 of the protocol.

A Steering Committee (SC) and an Independent Data Monitoring Committee (IDMC) will be commissioned for this study. The SC will oversee the study conduct and the IDMC will review safety data periodically. In addition, an interim analysis is planned when primary efficacy events have been observed in approximately 600 unique subjects. Depending on interim analysis results, the IDMC will make a recommendation whether the study will be continued as planned, or terminated early for overwhelming superiority, futility, or safety.

A diagram of the study design is provided below.



Note: 30 days maximum for a hospitalization for subjects enrolled in Japan

1.3. Statistical Hypotheses for Trial Objectives

The primary analysis endpoint for efficacy assessment (referred to as the primary efficacy endpoint hereafter) is the time from randomization to the first occurrence of ACM, MI, or stroke. The primary objective of the study will be addressed by comparing the distribution of the primary efficacy endpoint between the 2 treatment arms. The primary statistical hypotheses are:

Null hypothesis H_0: there is no difference between treatment arms in the distribution of the primary efficacy endpoint.

Alternative hypothesis H_A: the distribution of the primary efficacy endpoint is different between the 2 treatment arms.

Analyses endpoints and corresponding statistical hypotheses for secondary efficacy outcomes are defined similarly as above.

1.4. Sample Size Justification

This is an event driven study. The study was designed to observe occurrences of the primary efficacy event in 984 unique randomized subjects, on or before the GTED, to have 90% power to detect a 20% relative risk reduction (RRR, defined as one minus hazard ratio) in the composite of ACM, MI, or stroke at 5%, 2-sided significance level. A total of approximately 5000 subjects will be randomly assigned to either the rivaroxaban arm or the placebo arm in a 1:1 ratio.

The above sample size calculation was estimated based on the following additional assumptions:

- Event rate for the composite endpoint in the placebo arm: 19%/year
- Permanent premature treatment discontinuation rate: 10%/year
- Lost-to-follow-up: 1%/year
- Duration of enrollment period: 24 months (see Attachment 1 for distribution of enrollment by month)
- Duration of study (from first subject randomized to last EOS visit): 31 months

The assumptions of RRR and event rate were based on observations from HF subjects in the ATLAS ACS 2 TIMI 51 study. Observed data from HF subjects of ATLAS ACS 2 TIMI 51 study are summarized in the table below (Source: Adhoc analyses on file).

	Event Rate (%/patient-year)		Hazard Ratio (95% CI)
	Rivaroxaban 2.5 mg bid	Placebo	
HF Subjects	(N=562)	(N=558)	
ACM/MI/Stroke	9.43	16.64	0.57 (0.41, 0.79)
ACM	3.54	8.31	0.43 (0.26, 0.70)
MI	5.73	8.47	0.67 (0.44, 1.03)
Stroke	1.08	1.60	0.69 (0.26, 1.81)

Based on historical data, the mortality rate in this study was expected to be higher, and the rate of MI was expected to be lower, than the observed corresponding rates in ATLAS ACS 2 TIMI 51 study. After adjusting for differences in these event rates, an event rate of 19%/year for the composite endpoint in the placebo group was used in the sample size calculation. Event rates in the placebo group for the components ACM, MI, and stroke are expected to be approximately 14%, 4%, and 2% per year, respectively.

Although the RRR observed in ATLAS ACS 2 TIMI 51 study was 43% (95% CI [21%, 59%]), a RRR of 20% was assumed in the sample size calculation. The lower RRR close to the lower bound of the 95% CI was used to be conservative because the subject population in the current study was considered different from the HF subgroup in ATLAS ACS 2 TIMI 51.

The study is revised to observe occurrences of the primary efficacy event in 1,200 unique randomized subjects based on a simulation study performed on available trial data, i.e., longer enrollment period, lower than anticipated primary efficacy event rate, higher than anticipated study drug discontinuation rate:

- Event rate in the placebo arm: approximately 16%/year based on the observed combined event rate at 14%/year (assumed 20% of RRR remains)
- Early permanent study drug discontinuation rate: approximately 13%/year
- Duration of enrollment period: approximately 48 months (projected)

The simulation study estimated that the duration of study may extend to approximately 54 months based on the fact that the original goal of enrolling a total of 5,000 subjects remains

unchanged and all subjects would be observed for approximately 6 months after the last subject is enrolled.

With an effect size of 20% and overall α level of 5% (2-sided), 1,200 primary efficacy events will produce a statistical power of approximately 80%.

An interim analysis will be conducted when primary efficacy events are observed in approximately 600 unique subjects (see Section 3 for more details).

1.5. Randomization and Blinding

Central randomization will be implemented in this study. Subjects will be randomly assigned to 1 of 2 treatment groups in a 1:1 ratio, 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 country. The interactive web response system (IWRS) will assign a unique treatment code, which will dictate the treatment assignment and matching study drug kits for the subject. Study subjects, investigators, the SC members, and the sponsor study team members are blinded to treatment assignment until the completion of the study and data unblinding. The IDMC, whose primary responsibility is to ensure subject safety by monitoring data periodically, will have access to unblinded data. Measures and procedures to protect against unblinding of treatment assignment and study integrity, including data flow and personnel, who have access to unblinded data, are specified in the IDMC charter.

While all subjects will have D-dimer data collected before the first dose of study drug, the collection of D-dimer data at week 4 will be limited to a randomly selected subset of 10% of subjects within each country. This selection will be done through IWRS and the selection will be independent of treatment assignment. Complete D-dimer data will not be available until after the database lock.

2. GENERAL ANALYSIS DEFINITIONS

2.1. Subject-Level Trial Milestone Dates

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

Global treatment end date (GTED): The date when 1,200 primary efficacy events are predicted to have occurred. The sponsor will notify all investigative sites of the GTED. Subjects will stop their study drug after taking the second and final dose of study drug on this date. Because GTED is defined based on site local time, the date may be off by 1 day when converted to a central time.

First dose date: the date on which the first dose of study drug is taken by the subject. If missing or incomplete, the date is set to the earliest logically possible date on or after randomization.

Last dose date: the date on which the last dose is taken by the subject. If the day of last dose study drug taken cannot be determined, the last dose day will be assumed to be 32*K days from

the last visit at which study drug is dispensed (*K* is the number of bottles dispensed at that visit), capped by GTED, Early Permanent Study Drug Discontinuation visit if applicable, last contact day, and the upper bound of the logically possible range if partial date is provided, whichever occurs earliest, and no earlier than the first dose date. The number 32 is selected in the imputation because a bottle contains 64 tablets, covering 32 days of study drug therapy.

Last contact date (also referred to as trial reference end date): the date of the last trial-related procedure. For survival subjects, it is defined as the maximum of

- Date of last office visit (scheduled or unscheduled visit)
- Date of the last follow-up contact (including last date on subject survival status recorded in eCRF)
- Date of the last known outcome event (OE) or adverse event (AE) status or lab sample collection reported on the OE or AE or lab electronic case report from (eCRF) pages, respectively

If a subject completes the randomization visit but no further visits and no other data is recorded on the eCRF, the randomization date will be considered as the last contact date.

The last contact date will be capped by the database lock date. For subjects who died during the study and the death date is missing, the trial reference end date is defined as the death date, if known, or the database lock date, whichever occurs earlier.

2.2. Pooling Algorithm for Analysis Centers

No pooling algorithm for analysis centers will be used. The primary efficacy analysis will be stratified by region defined in Section 2.4.

2.3. Analysis Sets

The definition of an analysis data set consists of these 2 elements: 1) *subject population*, which specifies which subjects will be included in an analysis; and 2) *observation period*, which specifies the time window within which data will be included in an analysis. Key subject populations and observation periods are defined below.

Subject Populations

Intention-to-treat (ITT) population: This subject population consists of all randomized unique subjects who have a signed valid informed consent.

Per-Protocol (PP) population: This population is a subset of the ITT population, consisting of subjects who do not have a major protocol deviation in the following categories (see FRM-15949 for the specific major protocol deviations that are exclusions from the PP population and identification/review process):

- Entered but did not satisfy criteria
- Received a disallowed concomitant treatment

- Received wrong treatment or incorrect dose (includes subjects with treatment compliance rate < 75% calculated according to a formula described in Section 4.3)
- Developed withdrawal criteria but not withdrawn
- Other (as appropriate)

If appropriate, additional ad hoc exclusion criteria for the PP population may be applied. Subjects to be excluded from the PP population will be determined and documented prior to database lock, except for those who are excluded from PP due to taking the wrong treatment . Because unblinded data are needed to determine whether a subject has taken incorrect study drug, subjects excluded from the PP population due to taking any incorrect study drug will be determined after data unblinding, according to the pre-specified rule.

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

Observation periods

Up–to-GTED: this observation period includes all data from randomization to the GTED, inclusively. For time to event analyses, subjects who do not have events to be analyzed on or before the GTED will be censored on the GTED or the last contact date whichever occurs first.

On-treatment: this observation period includes data from the first dose of study drug to 2 days after the last dose of the study drug, inclusively. For time to survival event analyses, subjects who do not have events to be analyzed for this period will be censored on the last day of the period (last dose + 2) or the last contact date whichever occurs first.

Post-randomization: this observation period includes all data from randomization to the last contact day. Last contact day is defined in Section 2.1. For time to event analyses, subjects who do not have events to be analyzed on or before the last contact day will be censored on the last contact day.

Post-first dose of study drug: This observation period starts from the day of the first dose of study drug and ends on the day of last contact, inclusively. Because subjects will be asked to take their first dose of study drug on the randomization day, while they are still in investigators' offices, for most of subjects, this observation period will be identical to post-randomization observation period.

Off-treatment: this observation period includes data from 3 days after the last dose of the study drug to the last contact date, inclusively.

Day 3 to day 30 after the last dose: the 3rd day after the last dose of study drug to the 30th day after the last dose of study drug or the last contact day, whichever is earlier.

2.3.1. Efficacy Analysis Set

2.3.1.1. Main Efficacy Analysis Set

The main efficacy analysis set for all efficacy endpoints (primary efficacy, secondary and other efficacy analyses) is the analysis set defined by up-to-GTED observation period and the ITT subject population. Subjects will be analyzed according to the treatment group they are assigned, irrespective of the actual treatment received.

2.3.1.2. Sensitivity Efficacy Analysis Sets

The following analysis sets will be used to assess the robustness of the primary and secondary efficacy analyses.

- Analysis set defined by the on-treatment observation period and PP population.
- Analysis set defined by the on-treatment observation period and the safety subject population. Subjects will be analyzed according to the treatment group they are assigned, irrespective of the actual treatment received.
- Analysis set defined by the post-randomization observation period and the ITT subject population.

2.3.1.3. Additional Analysis Sets

The additional analysis sets are defined by the combination of observation periods:

- Off-treatment
- Day 3 to day 30 after the last dose

and, the subpopulations of:

- Subjects who permanently discontinue study drug for any reason prior to GTED
- Subjects who complete the double-blind treatment period and have the EOS visit (Office/Phone)

These analysis sets will be used mainly for the summary statistics for efficacy endpoints and are further described in Section 5.5.

2.3.2. Safety Analysis Set

The main safety analysis set is defined by the On-treatment observation period and the safety subject population. The sensitivity safety analysis set is defined by the post-first dose of study drug observation period and the safety subject population. Subjects will be analyzed according to the treatment assigned.

Bleeding events and adverse events will be analyzed and summarized based on both the main and sensitivity safety analysis sets.

In addition, the same additional analysis sets defined in Section 2.3.1.3 will also be used for the summaries of safety events. Details are described in Section 6.1.

2.4. Definition of Subgroups

Homogeneity of treatment effects, both in RRR and direction will be assessed in the following subgroups classified at baseline.

- Age (< 65 vs ≥ 65; < 75 vs ≥ 75 years)
- Sex (men vs women)
- LVEF (< 30% vs ≥ 30%; ≤ median vs > median)
- Estimated glomerular filtration rate (eGFR) (Modification of Diet in Renal Disease formula value <30, 30 to <60, 60 to <90, ≥90 mL/min/1.73 m^2)
- Baseline D-dimer level by quartile
- Baseline BNP or NT-proBNP (≤ median vs > median)
- History of diabetes (yes vs no)
- History of stroke (yes vs no)
- History of MI (yes vs no)
- Hypertension (yes vs no)
- Body Mass Index (<25, 25 to <30, ≥30 kg/m^2)
- Baseline digoxin use (yes vs no)
- Baseline β-blocker use (yes vs no)
- Baseline aldosterone antagonists (yes vs no)
- Baseline angiotensin-converting enzyme inhibitor (ACEI) or angiotensin receptor blocker (ARB) (yes vs no)
- Baseline aspirin use (yes vs no)
- Baseline aspirin only vs dual antiplatelet use
- Baseline thienopyridine use (yes vs no)
- NYHA (Class II, III, IV)
- Fragile subjects (yes vs no; fragile subjects are those with age > 75 years or weight ≤ 50 kg or baseline eGFR < 50 mL/min)
- Race (White vs others; White, Asian, and others)
- Region (Eastern Europe [Bulgaria, Czech Republic, Hungary, Latvia, Lithuania, Poland, Romania, Russia, Slovakia, Turkey, and Ukraine], Western Europe and South Africa [Denmark, France, Germany, Greece, Italy, Netherlands, Portugal, South Africa, Spain, Sweden, United Kingdom], North America [Canada and USA], Latin America [Argentina, Brazil, and Mexico], and Asia Pacific [Australia, China, Japan, Malaysia, and South Korea].

In addition, the number of subjects with efficacy endpoints including components of composite efficacy endpoints; the number of subjects with principal safety endpoint and its components of fatal bleeding (defined in Section 6.1) and bleeding into a critical space with a potential for permanent disability; and the number of subjects with other bleeding outcomes of ISTH major bleeding events and bleeding requiring hospitalization will be summarized for safety subjects in the following subpopulations:

- Subjects who permanently discontinue study drug for any reason prior to GTED
- Subjects who complete the double-blind treatment period and have the EOS visit (Office/Phone)

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

An interim analysis will be conducted when primary efficacy events with adequate supporting documents have been observed in approximately 600 unique subjects. The analysis will be carried out by an independent statistical group that supports the IDMC (referred to as statistical supporting group [SSG]). The IDMC will be responsible for the review of the interim analysis results. If early stopping criteria are met, the IDMC may recommend stopping the trial early for overwhelming treatment benefits or for futility/feasibility. Early stopping guidelines will be established and specified in the IDMC charter, which will be finalized prior to the first safety data review and therefore before the interim analysis. In addition, the IDMC will review safety data periodically. The frequency of safety reviews will also be specified in the IDMC charter.

In the interim analysis, a log-rank test, stratified by region, will be employed to compare the distribution of time to the first occurrence of the composite primary efficacy endpoint between the 2 treatment groups. The Lan-DeMets α spending function approach with O'Brien-Fleming type of boundaries will be used in the interim analysis for the primary efficacy endpoint only. The α spent in the interim analysis is expected to be 0.003 with a corresponding critical value of 2.936 (East®, Version 5.3, Cytel Inc.), if the interim analysis involves 600 events. If the study continues beyond the planned interim look, the critical value for the final analysis is 1.969, the corresponding α is 0.04895. The actual α spent and corresponding critical values may be slightly different from aforementioned numbers, depending on the actual number of events included in the interim analysis. If the test statistic for the primary efficacy endpoint crosses the stopping boundaries either at the interim analysis or at the final analysis, secondary endpoints will be tested in a fixed sequence as shown in Section 5.3.1, at a significance level corresponding to the Pocock-type boundary.

If the conditional power under the alternative hypothesis (i.e., RRR=20%) is 10% or lower, the IDMC may recommend stopping the study for futility.

In order to maintain blinding of randomized treatment assignments and study integrity, measures will be taken to ensure the separation between the study team and unblinded data. Data flow and parties who have access to unblinded data will be specified in the IDMC charter.

4. SUBJECT INFORMATION

4.1. Demographics and Baseline Characteristics

Descriptive statistics (mean, standard deviation, median, inter-quartile range, and range for continuous data, and frequency and percentage for categorical data) by treatment group will be provided for the following baseline demographics and disease characteristics data

- Age
- Sex
- Race
- Region
- Height
- Weight
- eGFR
- eGFR Category [< 30, ≥ 30 - < 60, ≥ 60 - < 90, ≥ 90]
- D-dimer
- Duration between discharge from index hospitalization to randomization
- NYHA classification
- LVEF at baseline (or last measurement of LVEF prior randomization)
- History of MI
- History of stroke
- History of PCI
- Cardiac resynchronization therapy or implantable defibrillator or combination
- Baseline digoxin use
- Baseline β-blocker use
- Baseline aldosterone antagonist use
- Baseline ACEI use
- Baseline ARB use
    - Baseline ARNI use
- Baseline aspirin use
- Baseline dual antiplatelet use
- Baseline thienopyridine use
- Baseline BNP or NT-proBNP

Additionally, evidence of coronary artery disease will be summarized.

4.2. Disposition Information

Unless otherwise stated, calculation of durations specified hereafter includes both the start and end days.

The number of subjects who permanently discontinue study drug and reasons for discontinuation will be summarized by treatment group. The distribution of time to permanent discontinuation of study drug, defined as the duration between the first dose and the last dose, will be shown by a modified version of Kaplan-Meier plot, where one minus the survival function will be shown.

A similar method will be used to summarize discontinuation from the study.

The primary efficacy analysis of this study will be addressed by data up to GTED. Therefore, data from subjects who are followed up until GTED are considered completers. Subjects who discontinue from the study prior to GTED without any primary efficacy event data between the last contact date and GTED are considered missing for analysis purposes. The fraction of follow-up time missing, defined as the ratio of the duration between the day next to the last contact date and GTED and the duration between the randomization date and GTED, will be summarized. The summary will be further broken down into missing data due to administrative reasons and missing data due to non-administrative reasons. Discontinuation reasons, such as lost-to-follow-up and consent withdrawal, are considered non-administrative.

A diagram similar to CONSORT Statement 2010 Flow Diagram (Schulz et al, 2010) will be provided. A template of the diagram can be found in Attachment 2.

4.3. Treatment Compliance

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

$$\text{Compliance rate (\%)} = \frac{100 * \text{Number of days of taking study drug}}{\text{Total treatment duration in days (excluding intended treatment interruptions)}}$$

where the number of days taking study drug is defined as the total treatment duration (from the first dose to the last dose) excluding all interruptions regardless of intentional or not.

Adherence to study regimen for an individual subject is considered acceptable in this study if the compliance rate is 75% or higher.

For the purpose of compliance rate calculation, interruptions due to AE, outcome events, bleeding events, prohibited medication, or surgical/invasive procedures are considered intentional interruptions. Interruptions due to other reasons are considered non-intentional. Missing doses during intentional interruption periods will not be counted in the compliance rate calculation. If the dose date prior to a treatment interruption is missing, the same imputation rule for imputing the last dose date descripted in Section 2.1 will be used to impute the dose date.

The median, the inter-quartile range, and the range of compliance rate will be summarized by treatment. In addition, if information is sufficient, the compliance rate in each treatment group will be depicted by a cumulative distribution function (CDF).

The proportion of subjects with acceptable treatment adherence as defined above will be provided.

4.4. Extent of Exposure

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

The mean, standard deviation, median, the inter-quartile range, and the range of duration of treatment exposure will be summarized by treatment. In addition, the duration in each treatment group will be depicted by a CDF.

The duration will also be summarized by the following categories:

- ≤ 90 days
- 91 – 180 days
- 181 – 360 days
- 361 - 540 days
- 541 – 720 days
- 721 - 900 days
- 901 – 1080 days
- 1081 – 1260 days
- > 1260 days

4.5. Protocol Deviations

Major protocol deviations will be summarized by treatment group for the all randomized data set. The categories of protocol deviations will include the following:

- Entered but did not satisfy criteria
- Received a disallowed concomitant treatment (Attachment 3)
- Received wrong treatment or incorrect dose (includes subjects with treatment compliance rate < 75% calculated according to a formula described in Section 4.3)
- Developed withdrawal criteria but not withdrawn
- Other (as appropriate)

4.6. Prior and Concomitant Medications

This study collects limited prior and/or concomitant medications, such as medications related to subject's HF and CAD. Medication will be classified using the Anatomical Therapeutic Chemical (ATC) classification system in WHO drug dictionary. The prevalence of medication use by ATC term will be summarized by treatment group and by study phase (prior to randomization, during the study drug therapy period, and after last dose of study drug [post-treatment]). A medication prior to randomization is considered as a prior medication. A medication with a use period overlapping with the study drug therapy period 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 can be categorized into multiple study phases if its use spans more than one study phase. In such a case, the medication use will be included in more than one summary.

5. EFFICACY

5.1. Analysis Specifications

Only outcome events with adequate supporting documentation as specified in the protocol will be included in the main analyses described below. All events reported by investigators, regardless of adequacy of supporting documentation, will only be included as part of sensitivity analyses. The adequacy of documentation will be determined prior to data unblinding.

If it is determined prior to data unblinding that fraud or misconduct occurred at a clinical study site during the conduct of the study, all subjects at that site will be excluded from primary efficacy analysis. Other findings of noncompliance with Good Clinical Practices (GCP) will be assessed to determine whether subjects affected will need to be removed from analyses. Subjects to be excluded from analyses due to fraud or misconduct of a clinical study site will be documented prior to data unblinding.

5.1.1. Level of Significance

The family-wise type I error rate for the primary efficacy endpoint will be controlled at the 5% level (2-sided). This is achieved by using the Lan-DeMets α spending function approach. If superiority is established for the primary efficacy endpoint, secondary endpoints will be tested in a fixed sequence, each at a significance level corresponding to the Pocock-type boundary. If an individual test during any step is not statistically significant, further treatment comparison may continue (i.e., reporting of p-values) but significance will not be claimed.

Confidence intervals (CIs) are calculated at the 95% nominal confidence level. In addition, repeated CI (RCI), which adjusts for the interim analysis, will also be provided in the Clinical Study Report (CSR).

5.1.2. Data Handling Rules

If the date of an outcome event is missing or partially missing, the date at the middle of the logically possible range will be imputed as the event date. If there is a tie between 2 such dates, (i.e., the number of days within the possible range is an even number), the later date will be used.

In the case that it cannot be determined if an event occurs before or after randomization, the event is always considered as a post-randomization event and the randomization date is used as the start of the logically possible range. On the other hand, in the case that it cannot be determined if an event occurs before or after GTED, the occurrence date is always considered not later than the GTED. In this case the GTED is the end of the logically possible range.

5.2. Primary Efficacy Endpoint(s)

5.2.1. Definition

The primary efficacy endpoint is the time from randomization to the first occurrence of ACM, MI, or stroke. Subjects who do not have any of these outcome events will be censored at the end of the analysis observation period as defined in Section 2.3.1.

5.2.2. Analysis Methods

The primary statistical null hypothesis is that there is no difference between treatment groups in distribution of the time from randomization to the first occurrence of ACM, MI, or stroke, and the alternative hypothesis is that there is a difference in the distribution between treatment groups.

The primary statistical hypothesis will be tested using a log-rank test, stratified by region. The primary analysis will be based on the analysis set defined by ITT subject population and the up-to-GTED observation period (up to the cut-off date to be specified for the interim analysis, see details below). Subjects will be analyzed according to the treatment group to which they are randomized, irrespective of actual treatment received. The overall α level is 5%, 2-sided. In addition to the final analysis, the primary statistical hypothesis will be tested in an interim analysis when primary efficacy events are observed in approximately 600 unique subjects. A Lan-DeMets α spending function with O'Brien-Fleming type of boundaries will be used to preserve the overall type I error rate. If the log-rank test statistic crosses the stopping boundaries, and the log-rank statistic is less than 0, (i.e., the observed number of events in the rivaroxaban group is less than the expected number of events under the null hypothesis), it will be concluded that the study has demonstrated that rivaroxaban is superior to placebo in reducing the risk of the composite of ACM, MI, and stroke in the population studied.

It is expected that risk of death will be the highest immediately following the index event or following subsequent episodes of re-hospitalization for heart failure. Therefore, the cumulative event rate derived from the Kaplan-Meier estimate will be displayed graphically to evaluate the timing of event occurrence and the consistency of the treatment effect over time.

The magnitude of the RRR will be estimated using a Cox proportional hazards model, stratified by region, with treatment (as randomized by IWRS) as the only covariate. The point estimate and corresponding 95% CI, which will be calculated using 95% as the nominal significance level, for the hazard ratio (rivaroxaban to placebo) will be reported. In addition, a RCI will be calculated based on the procedure described in Jennison and Turnbull (2000), where significance level at the interim analysis and the final analysis will be approximately 99.7% and 95.2%, respectively. The plausibility of the proportional hazards assumption will be assessed by visually comparing

the plot of the log of the cumulative hazard function between treatments, and formally tested by adding a treatment by logarithm-transformed time interaction into the Cox model. A p-value greater than 0.05 for the interaction term will be interpreted as no statistical evidence against the proportional hazard assumption. If the p-value is 0.05 or less, post-hoc analyses will be conducted to estimate the magnitude of treatment effect by time. These post-hoc analyses include, but are not limited to, adding a function of time, such as time interval indicators, by treatment interaction as covariates to the Cox model. Treatment effects for different time intervals, such as < 12 months versus > 12 months, etc., will be reported as appropriate.

Sensitivity analyses will also be conducted to assess the robustness of the primary efficacy analysis. These include, but are not limited to, a log-rank test stratified by country and an un-stratified log-rank test, repeating analyses mentioned above for the analysis set defined by the up-to-GTED observation period and the ITT subject population; for the analysis set defined by the on-treatment observation period and the per-protocol subject population; for the analysis set defined on-treatment observation period and the safety subject population; and for the analysis set defined by the post-randomization observation period and ITT subject population. The primary efficacy endpoint will also be analyzed for the ITT subjects who enrolled under the original protocol and who enrolled under the amended protocols for the Up-to-GTED observation period.

The percentage of events reported by investigators with adequate protocol-specified documentation will be reported.

The RRR for the components of the primary efficacy endpoint will be evaluated using a Cox proportional hazards model as described above.

Extensive efforts will be made to collect complete data for all subjects randomized in this study. Subjects will be followed up to the end of the study, and all required data will be collected, regardless of subjects' compliance with study drug regimen or visit schedule. For subjects who are lost to follow up or who withdraw consent, efforts will be made to obtain their vital status at the end of study from permitted sources.

Although extensive effort will be made to reduce the number of subjects with missing follow-up, it is expected that there will be missing vital status and event information in some subjects. Sensitivity analyses will be performed to evaluate the potential impact of missing data on analysis results and robustness of study conclusions. Distributions of baseline demographics and other characteristics will be compared between subjects who complete the study and subjects who do not complete the study, accounting for treatment and regardless of treatment, to evaluate the plausibility of the non-informative assumption used in the analysis. In addition, reasons of early discontinuation from the study will be summarized by treatment group.

The potential impact of missing data on the primary efficacy analysis will be evaluated by imputation, where events will be imputed in missing follow-up periods. The imputation will be done based on various assumptions for unobservable event rates in missing follow-up periods, which may depend on treatment groups and/or reasons of early discontinuation. Additional

details of these sensitivity analyses, including specifications for imputation, will be provided in a separate document that will be finalized prior to the database lock.

Homogeneity of treatment effects, both in the magnitude and direction of RRR, across subgroups will be assessed by subgroup analyses. Subgroups to be included in the subgroup analyses are defined in Section 2.4.

Interactions between treatment and baseline covariates with continuous values, such as age and D-dimer, will also be evaluated on a continuous scale in addition to categorical scales.

For subgroups specified in Section 2.4, annualized event rates will be provided by treatment and by subgroup. If there are at least 10 events in each of the cells defined by combinations of treatment and subgroup, statistical testing will be done to test the homogeneity of the treatment effect hypothesis. Statistical testing will be conducted for each set of subgroups individually, by adding covariate(s) for the subgroup and the corresponding treatment-by-subgroup interaction in the Cox proportional hazards model specified above, one at a time, except for the assessment of a region effect. An un-stratified Cox model will be used when assessing the homogeneity of treatment effect across regions. For subgroups with natural ordering, such as subgroups defined by eGFR, Body Mass Index, and NYHA classification, an ordinal covariate will be created to indicate subgroups. If a non-monotonic trend is observed, additional subgroup analyses based on non-ordinal categories may be further explored. For other subgroups without natural ordering, such as subgroups defined by race or region, a set of indicator covariates will be created to indicate subgroups. If the p-value for the interaction is greater than 0.05, the test result will be interpreted as a lack of statistical evidence for a difference in the treatment effect between the subgroups. Otherwise, a Gail-Simon test will be performed to assess if the interaction is also qualitative. If the p-value from the Gail-Simon test is greater than 0.05, the result will be interpreted as a difference in the size of treatment effect between the subgroups with a lack of statistical evidence for the difference in the direction of treatment effect (i.e., only quantitative but not qualitative interaction). If the p-value is less than or equal to 0.05, it will be interpreted as a directional difference in treatment effect between subgroups (i.e., qualitative interaction). Given the large number of subgroup analyses explored without multiplicity adjustment, observing p-values less than 0.05 due to chance alone is expected. Therefore, results of the subgroup analyses will be interpreted with caution.

The above univariate subgroup analyses will be supplemented by a multivariate analysis, which will assess the effects of baseline demographics and disease characteristics collectively using an un-stratified Cox proportional hazards model that contains main effects of treatment and subgroups, and interactions between treatment and subgroup (2-way interactions only). This approach adjusts for correlations between variables and evaluates the net effects of subgroups. A backward selection procedure will be employed. In the backward selection procedure, the variable with the largest p-value will be eliminated from the model one at a time. A main effect cannot be removed from the model before the corresponding interaction with treatment is removed. The backward selection procedure ends when no more covariates can be removed from the model with the exit p-value threshold of 0.05.

5.3. Secondary Efficacy Endpoints

5.3.1. Definition

There are 4 secondary efficacy endpoints specified in this study:

- Time to the first occurrence of CV death or re-hospitalization for worsening of HF
- Time to CV death
- Time to re-hospitalization for worsening of HF
- Time to re-hospitalization for CV events

5.3.2. Analysis Methods

If superiority of rivaroxaban over placebo with respect to the primary efficacy endpoint is established, the secondary endpoints will be tested in the sequence specified above. The determination of the order is based on clinical relevance and power consideration. Each secondary endpoint will be tested using the same analysis methods that are described for the primary efficacy endpoint. The significance level of each individual test corresponds to the Pocock-type boundary, irrespective of whether the study is completed as planned or stopped early at the interim analysis. If p-value for a hypothesis is greater than that significance level, all remaining hypotheses after this hypothesis in the fixed sequence will not be rejected.

A supporting analysis of time to the first occurrence of ACM or re-hospitalization for worsening of HF will also be performed using the same analysis method as for the primary efficacy endpoint.

As supplemental analyses, the effect size of rivaroxaban versus placebo in terms of re-hospitalization for worsening of HF and for CV events will also be estimated using parametric, semi-parametric, and non-parametric approaches. In contrast to time to the first occurrence used in the analyses mentioned above, all occurrences will be included in these supplemental analyses. A parametric frailty model will be used in the parametric approach. The frailty model assumes that events follow a renewal process with a time-independent hazard that varies from subject to subject, and the distribution of the subject-dependent hazard follows a Gamma distribution, which is treatment specific. The resulting analysis is a generalized linear regression model based on negative binomial distribution, where the covariates are duration of risk exposure, treatment, and the interaction between them. The semi-parametric approach is the Andersen-Gill model (Andersen and Gill, 1982) with a robust estimate for variance that takes intra-subject correlation into account (Lin and Wei, 1989). The non-parametric analysis is based on the unmatched win-ratio method proposed by Pocock et al (2012).

5.4. Other Efficacy Variable(s)

5.4.1. Definition

Analysis of the following efficacy endpoints are considered exploratory:

- Time to symptomatic deep vein thrombosis (DVT)

- Time to symptomatic pulmonary embolism (PE)

5.4.2. Analysis Methods

The number of symptomatic DVT and PE events is expected to be small. Due to the exploratory nature of the analysis and consideration of the type I and type II errors, no statistical testing will be conducted for these 2 exploratory endpoints. However, the size of the treatment effects, including corresponding 95% CI, of time to DVT and time to PE will be estimated using a Cox proportional hazards model with treatment as the only covariate. Because of the small number of events expected, the Cox model is not stratified.

In order to understand the impact of strokes, the modified Rankin score will be collected within 6-18 weeks following a stroke or at the end of study, whichever occurs first. The distribution, (i.e., frequency and percentage), of the modified Rankin score will be tabulated for stroke events.

5.5. Additional Efficacy Analyses

For all primary, secondary and other efficacy outcome events discussed above, the number of subjects with these events will be tabulated for the following analysis sets:

- Off-treatment observation period for subjects who permanently discontinue study drug for any reason prior to GTED
- Off-treatment observation period for subjects who complete the double-blind treatment period and have the EOS visit (Office/Phone)
- Off-treatment observation period for the Safety population.
- Day 3 to Day 30 after the last dose observation period for subjects who permanently discontinue study drug for any reason prior to GTED
- Day 3 to Day 30 after the last dose observation period for subjects who complete the double-blind treatment period and have the EOS visit (Office/Phone)
- Day 3 to Day 30 after the last dose observation period for the Safety population.

6. SAFETY

6.1. Bleeding Events

The principal safety endpoint for this study is the composite of fatal bleeding or bleeding into a critical space with a potential for permanent disability. A fatal bleeding event is one in which the subject dies within 7 days of a bleeding event requiring hospitalization or ISTH major bleeding. Fatal bleeding events will include those that meet the criteria in the following three exclusive categories:

- Category 1: Any ISTH major bleeding event considered to be the primary cause of death by the investigator (i.e., Question 5 on the BLEED page of the eCRF is answered 'Yes'), or

- Category 2: Any ISTH major bleeding event not considered to be the primary cause of death by the investigator (i.e., Question 5 on the BLEED page of the eCRF is answered as 'No') but results in death within 7 days, or
- Category 3: Any bleeding event that results in a hospital stay and death within 7 days (i.e., death and either Question 6 or 8 on the BLEED page of the eCRF is answered 'Yes')

The following organs are considered critical:

- Intracranial
  - Intraparenchymal
  - Intraventricular
  - Subdural Hemorrhage
  - Subdural
  - Epidural
- Intraspinal
- Intraocular—other than sub-conjunctival
  - Vitreous
  - Retinal
- Pericardial
- Intra-articular
- Intramuscular with compartment syndrome
- Retroperitoneal

Similar to the analyses planned for the primary efficacy endpoint as described above, time to the first occurrence of the principal safety endpoint will be compared using a log-rank test and a Cox proportional hazards ratio model with treatment as the only covariate. Both the log-rank test and Cox proportional hazards models are stratified by region. A plot of cumulative event rate derived by Kaplan-Meier estimate will be provided to show event rate and treatment effect by time. The analyses will be conducted based on analysis set defined by the on-treatment observation period and the safety subject population. Subjects will be analyzed according to study drug assigned. Significance level is always at the nominal level of 0.05 for safety analyses. There will be no adjustment for multiplicity. The same analysis will be repeated for the post-first dose observation period and the safety subject population.

Time to the first occurrence of following bleeding events will be analyzed using the same method.

- Bleeding requiring hospitalization
- ISTH major bleeding

The number of subjects with the above safety outcomes will be tabulated for the following observation periods and analysis sets:

- Off-treatment observation period for subjects who permanently discontinue study drug for any reason prior to GTED

- Off-treatment observation period for subjects who complete the double-blind treatment period and have the EOS visit (Office/Phone)
- Off-treatment observation period for the Safety population.
- Day 3 to Day 30 after the last dose observation period for subjects who permanently discontinue study drug for any reason prior to GTED
- Day 3 to Day 30 after the last dose observation period for subjects who complete the double-blind treatment period and are have the EOS visit (Office/Phone)
- Day 3 to Day 30 after the last dose observation period for the Safety population.

The number of the subjects with other reported bleeding events will be summarized by treatment group and bleeding site. For Japan subjects only, summary statistics will be provided for non-major clinically relevant bleeding events and minimal bleeding events.

6.2. Other Adverse Events

Because the safety profile of rivaroxaban has been well established in previous large and extensive trials, this study will collect limited AE data (see Section 12 of the study protocol for AE data to be collected). 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). For each MedDRA preferred term, the percentage of subjects who experience at least 1 occurrence of the given event will be summarized by treatment group. The difference between treatments in the percentage will be provided. If an AE is reported by 4 or more subjects in the entire safety population, a 95% CI for the difference will also be provided. The 95% CI is calculated based on the normal approximation of binomial data without continuity correction. This is corresponding to the approach for Tier 2 events (signal detection among common events) proposed by Crowe et al (2009). Additional summaries, listings, or subject narratives may be provided, as appropriate.

AEs reported by investigators will be summarized based on the analysis set defined by the on-treatment observation period and separately by the post-first dose of study drug observation period and the safety subject population.

Based on specific safety reporting requirements from the Pharmaceuticals and Medical Devices Agency, AE data from Japan may not be directly comparable with those collected from non-Japan sites. Therefore, the AE data from Japan sites may be summarized separately from AE data from non-Japan sites in the respective reported safety tables.

6.3. Clinical Laboratory Tests

Because the safety profile of rivaroxaban has been well established in previous trials, this study will not collect laboratory data during the study other than D-dimer at Week 4 (10% of subjects) and hemoglobin at Week 12 (up to Week 24). Complete D-dimer data will not be available until after the database lock. There is no plan to analyze laboratory data other than a descriptive summary of the hemoglobin and D-dimer data. Local lab data may be discussed in subject narratives as appropriate.

Change from baseline in Week 4 D-dimer, in the 10% of subjects from whom Week 4 D-dimer values will be collected, will be summarized by treatment group. A similar approach will be used to summarize the hemoglobin data. Mean, standard deviation, median, inter-quartile range, and range will be reported.

6.4. Vital Signs and Physical Examination Findings

Targeted medical history and vital signs are obtained at baseline only and summarized as baseline characteristics Physical examinations are not required as part of study procedures. Available data may be discussed in subject narratives.

7. Benefit/Risk Assessment

The section below 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 population and observation period, and 4) reporting format of the results.

The benefit-risk analyses described here are primarily intended to structure an integrated evaluation of the key benefits and risks in the study. They are complementary to other efficacy and 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 absolute risk differences. 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 or person-years). The event rate will be calculated based on the time-to-first-event using the following approaches:

- Person-year rate, expressed as number of events per 100 person-years exposure time
- Kaplan-Meier Product-Limit estimates of cumulative event rates at 1-year*

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.

*Because the duration of the trial is longer than anticipated, Kaplan-Meier Product-Limit estimates of cumulative event rates at 2-years and 3-years may be explored.

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. The primary efficacy outcome is the composite of ACM, MI, or stroke, while the principal safety outcome is the composite of fatal bleeding or bleeding into a critical space (intracranial, intraspinal [vitreous or retinal], pericardial, intraocular-other than sub-conjunctival, intra-articular, retroperitoneal, intramuscular with compartment syndrome, retroperitoneal) with a potential for permanent disability. Importantly, for the benefit-risk calculations, fatal bleeding and hemorrhagic stroke would be counted in both the benefit and risk categories (e.g., fatal bleeding will be part of ACM, while hemorrhagic stroke will be part of stroke). Therefore, to avoid double-counting an event as both a benefit and as a risk, the outcomes for the benefit-risk analysis will be the following:

Benefits

Composite of non-bleeding related death, non-fatal MI, or non-fatal ischemic stroke and each component separately:

- Non-bleeding related death
- Non-fatal MI
- Non-fatal ischemic stroke

As a supporting measure, the composite of non-bleeding related death, MI (fatal, non-fatal) or ischemic stroke (fatal, non-fatal) may also be evaluated.

Risks

Composite of fatal bleeding events, non-fatal intracranial hemorrhage, or other bleeding into a critical space with a potential for permanent disability and each component separately:

- Fatal bleeding events
- Non-fatal intracranial hemorrhage
- Non-intracranial hemorrhage, non-fatal bleeding into a critical space with a potential for permanent disability

The two composite endpoints listed above will serve as the primary comparison of benefits and risks (pairwise comparison), since those events are fatal or are likely to cause irreversible harm (Unger 2009, Beasley 2011) and arguably represent events of similar clinical importance. In addition, components of the composite endpoints and 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 Population and Observation Period

The primary analysis for benefit-risk evaluation will be based on the analysis set defined by the ITT subject population and the up-to-GTED observation period. However, additional analyses using analysis sets defined by other subject populations and/or observation periods as defined in the protocol or in this SAP may be performed as supporting analyses.

Note that 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.

Note that although primary composite endpoints and analysis set are mentioned above, assessment of benefit/risk balance is based on the totality of data. It is unlikely that a clear conclusion can be made based on a comparison of a pair of composite endpoints and/or data from a typical analysis data set. Those primary composite endpoints and primary analysis data set provide a starting point. Results of the analyses do not preclude additional benefit/risk assessment in other endpoints and/or analysis population, and they do not carry more weight than other additional analyses in the overall benefit/risk assessment.

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 (for the pairwise comparison)
- Forest plots for comparing key benefits and risks (as discussed above), as well as other main efficacy and safety outcomes measures

8. HEALTH ECONOMICS

Medical Resource Utilization data will be analyzed and reported in a separate document.

REFERENCES

1. Andersen, P.K., Gill, R.D. (1982). Cox regression model for counting processes: A large sample study. *Annals of Statistics*: 10, 1100–1120.
2. Beasley BN, Unger EF, Temple R (2011). Anticoagulant options ─ why the FDA approved a higher but not a lower dose of dabigatran. *N Engl J Med*;364(19):1788-90.
3. Crowe BJ, et al (2009). Recommendations for safety planning, data collection, evaluation and reporting during drug, biologic and vaccine development: a report of the safety planning, evaluation, and reporting team. *Clinical Trials*: 6, 430 – 440.
4. Gail, M. and Simon, R. (1985). Testing for Qualitative Interactions between Treatment Effects and Patient Subsets. *Biometrics*, Vol. 41, No. 2, 361-372
5. Glimm, E., Maurer, W., and Bretz, F. (2010). Hierarchical testing of multiple endpoints in group-sequential trials. *Statistics in Medicine*, 29 219-228
6. Jennison, C. and Turnbull, B. W. (2000) Group Sequential Tests with Applications to Clinical Trials, *Chapman and Hall/CRC*, 390 pages
7. Lin, D. Y. and Wei, L. J. (1989). The Robust Inference for the Proportional Hazards Model. *J of American Statistical Association:* 84, 1074–1078.
8. Pocock S. J., Ariti C. A., Collier T. J., and Wang D. (2012). The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *European Heart Journal:* 33, 176–182.
9. 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).
10. Tamhane, A.C, Mehta, C. R., Liu, L. (2010). Testing a Primary and a Secondary Endpoint in a Group Sequential Design. *Biometrics*, 66(4): 1174-1184
11. Unger EF (2009). Weighing benefits and risks ─ the FDA's review of prasugrel. *N Engl J Med*;361(10):942-5.

ATTACHMENTS

Attachment 1. The assumption of enrollment used in the original sample size calculation and the observed enrollment up to June 2016 on which the simulation study was based

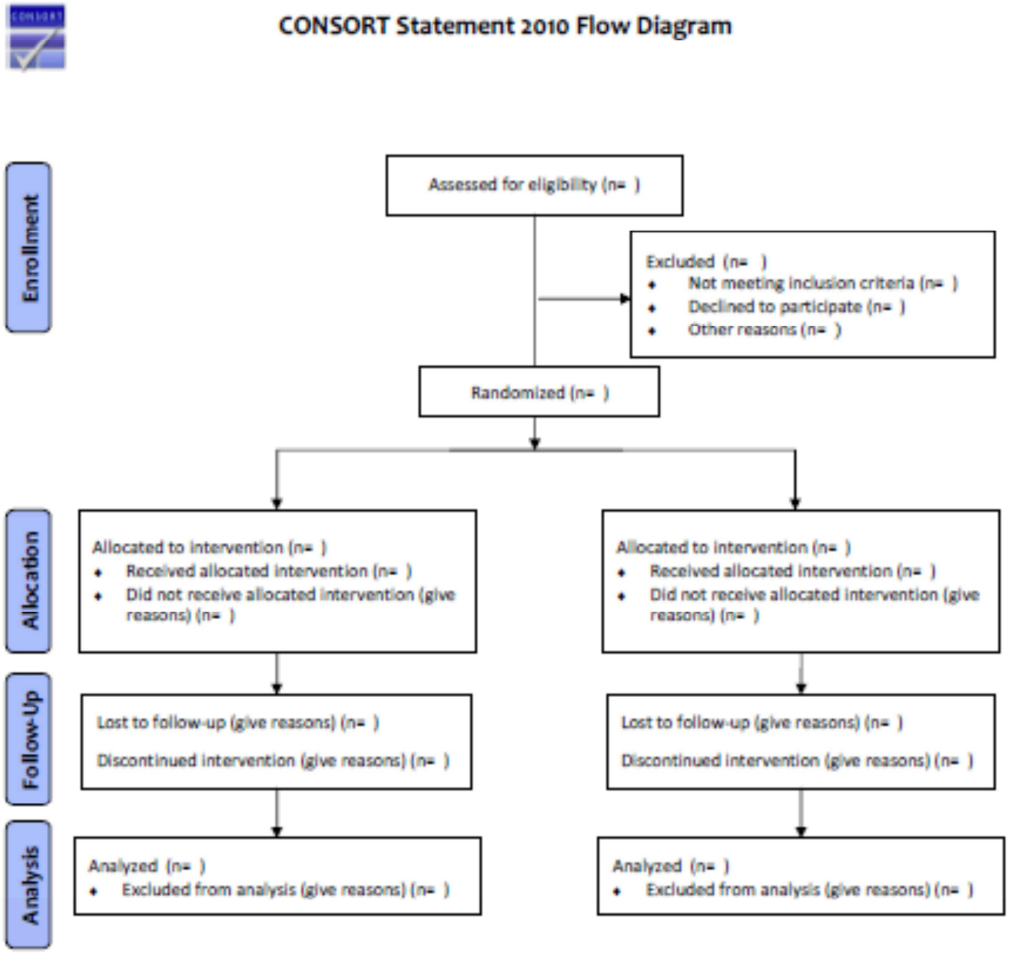
The table below specifies monthly enrollment assumed in the sample size calculation.

Month	Number of subjects enrolled in the month
Month 1	2
Month 2	5
Month 3	13
Month 4	25
Month 5	50
Month 6	85
Month 7	120
Month 8	160
Month 9	210
Month 10	270
Month 11 and afterward	320

The table below shows the observed subject enrollment up to June 2016

Calendar Month	Number of subject randomized
Sep-13	4
Oct-13	9
Nov-13	51
Dec-13	83
Jan-14	48
Feb-14	91
Mar-14	108
Apr-14	135
May-14	107
Jun-14	112
Jul-14	123
Aug-14	68
Sep-14	78
Oct-14	120
Nov-14	73
Dec-14	71
Jan-15	67
Feb-15	70
Mar-15	61
Apr-15	63
May-15	68
Jun-15	65
Jul-15	53
Aug-15	62
Sep-15	67
Oct-15	109
Nov-15	125
Dec-15	126
Jan-16	125
Feb-16	117
Mar-16	164
Apr-16	161
May-16	137
Jun-16	178

Attachment 2. Template of CONSORT Statement 2010 flow diagram



Attachment 3. Disallowed concomitant medications

The following medications are prohibited in this study:

- Prasugrel is prohibited in subjects who are ≥75 years old in age, or in subjects with prior TIA or stroke.
- Nonsteroidal anti-inflammatory agents (NSAIDs) may be used on a temporary basis, but should be avoided for chronic use during the study period.
- Strong inhibitors of cytochrome P450 3A4, such as but not limited to, ketoconazole, itraconazole, telithromycin, clarithromycin and voriconazole or protease inhibitors, are prohibited as concomitant therapy within 4 days before randomization, or during the study.
- Any drug which is contraindicated in patients with heart failure (e.g., cilostazol)
- Strong inducers of cytochrome P450 3A4, such as but not limited to, rifampin/rifampicin, rifabutin, rifapentin, phenytoin, phenobarbital, primidone, St. John's Wort, or carbamazepine are prohibited as concomitant therapy within 4 days before randomization, or during the study.
- Proton pump inhibitors may be used, however subjects receiving clopidogrel should not receive omeprazole or esomeprazole.