Statistical and Epidemiological Analysis Plan – Objective 1

c13963790-01

BI Trial No.:	1160.188
Title:	Cha**R**act**E**rization of patients following a**C**ute venous thr**O**mboembolism (**VTE**) and assessment of safety and effectiveness of dabigatran etexilate (DE) in the t**R**eatment and secondar**Y** prevention of acute deep vein thrombosis (**DVT**) and pulmonary embolism (**PE**) in comparison to vitamin K antagonist (VKA) in routine clinical practice - **RE-COVERY DVT/PE** Including protocol amendment 1 [c03118416-02]
Investigational Product:	Dabigatran etexilate
Responsible trial statistician(s):	Phone: Fax: Phone: Fax:
Responsible trial epidemiologist:	Phone: Fax:
Date of statistical and epidemiological analysis plan:	05 APR 2017 SIGNED
Version:	FINAL

1. TABLE OF CONTENTS

LIST OF TABLES

2. LIST OF ABBREVIATIONS

Term	Definition / description
ADR	Adverse drug reaction
AE	Adverse event
BI	Boehringer Ingelheim
b.i.d.	bis in die (two times a day)
BMI	Body Mass Index
CI	Confidence interval
DVT	Deep venous thrombosis
ECG	Electrocardiogram
eCRF	Electronic case report form
EDC	Electronic data capture
EoT	End-of-text
ICH	International Conference on Harmonisation
INR	International normalized ratio
IPV	Important protocol violation
ISPE	International Society for Pharmaceutical Engineering
LMWH	Low molecular weight heparin
LOS	Length of stay
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
MQRM	Medical quality review meeting
NIS	Non-interventional study
NOAC	Non-vitamin K anticoagulant
PACT-Q©	Perception of Anticoagulant Treatment Questionnaire
PE	Pulmonary embolism
PV	Protocol violation
Q1	Lower quartile
Q3	Upper quartile
SD	Standard deviation
SEAP	Statistical and epidemiological analysis plan
UFH	Unfractionated heparin

Term	Definition / description
VKA	Vitamin K antagonist
VTE	Venous thromboembolism
WHO	World Health Organization

3. INTRODUCTION

As per ICH E9 [1] and the ISPE Guidelines of Good Pharmacoepidemiology Practice [2], the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and to include detailed procedures for executing the statistical analysis of the data. This document describes the final analysis at the end of study for Objective 1 as well as the interim analyses which will be performed at the end of each of the first two enrolment periods for Objective 1.

This SEAP assumes familiarity with the non-interventional study (NIS) protocol including Protocol Amendments. In particular, the SEAP is based on the planned analysis specification as written in the NIS protocol Section 9 “Research Methods”. Therefore, SEAP readers may consult the NIS protocol for more background information on the study, e.g., on study objectives, study design and population, treatments, definition of measurements and variables, planning of sample size.

SAS® Version 9.3 or higher will be used for all analyses.

4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

It was originally planned in the protocol that approximately 2,000 patients would be enrolled into Objective 1 per year over three years for a total of 6,000 patients, and that the study would include an analysis of the patients from each enrollment year through two interim analyses. Subsequently, a decision was made by the study team to continue Objective 1 enrollment without pausing between years until the total target of 6,000 patients is reached. As a result, it is projected that the enrollment for Objective 1 will be completed in less than two years. In light of this change, it is now planned that only one interim analysis will be performed with data from approximately the first 3,000 patients.

5. OUTCOMES

Objective 1 of the RE-COVERY study is designed to characterize the deep vein thrombosis (DVT)/pulmonary embolism (PE) patient population including the initial acute event phase. All patients with a DVT and/or PE will be enrolled for cross-sectional characterization of the venous thromboembolism (VTE) patient population and descriptions of current treatment patterns, and stratified by geographical location.

The following research questions will be investigated:

- With the approval of dabigatran etexilate and other non-vitamin K anticoagulants (NOACs), what treatments are being administered for acute VTE in routine clinical practice in the different regions of the world?

- When presenting with an acute VTE event, what factors influence the choice of treatment for the event?

5.1 MAIN OUTCOMES

The following primary outcome measures will be analyzed on data that are collected at the time of the acute VTE event:

- Demographic information (age and gender)

- VTE event information (event type and treatment for event)

6. GENERAL ANALYSIS DEFINITIONS

6.1 EXPOSURE

For Objective 1, VTE treatment information will be collected at two time points, baseline and hospital discharge/14 days after baseline, due to the requirement of heparin pre-treatment for some of the studied oral anticoagulants.

On the eCRF page, five oral anticoagulants (vitamin K antagonist, dabigatran, edoxaban, rivaroxaban, and apixaban) are included as pre-specified options, together with three types of heparin pre-treatment (low molecular weight heparin, unfractionated heparin, and fondaparinux) and two types of non-oral therapies (systemic thrombolytic therapy and catheter directed thrombolytic therapy). Patients will be assigned to one of the five oral anticoagulant groups or an additional "Other" group (defined as patients with no oral anticoagulant prescribed), based on the following rules.

Table 6.1: 1 Oral anticoagulant treatment assignment

Treatment at baseline	Treatment at hospital discharge/14 days	Treatment assignment
Vitamin K antagonist[1]	Same as baseline	Vitamin K antagonist
Dabigatran[1]		Dabigatran
Edoxaban[1]		Edoxaban
Rivaroxaban		Rivaroxaban
Apixaban		Apixaban
Heparin pre-treatment without any oral anticoagulant and without non-oral/other therapy	Vitamin K antagonist	Vitamin K antagonist
	Dabigatran	Dabigatran
	Edoxaban	Edoxaban
	Rivaroxaban	Rivaroxaban
	Apixaban	Apixaban
Any of the following without oral anticoagulant: Systemic thrombolytic therapy Catheter directed thrombolytic Other	Same as baseline	Other

[1] Although heparin pre-treatment is required for vitamin K antagonist, dabigatran and edoxaban, these scenarios are included to account for the possibility that these treatments were entered on the baseline eCRF page by the sites due to the retrospective enrollment of patients.

In the event that patients with other combinations of treatment information at baseline and hospital discharge/14 days, such as those listed below, are identified, their treatment assignment should be discussed and determined on a case-by-case basis by the study team during the periodical medical quality review meetings (MQRMs). In the event that a decision

cannot be made due to contradictory information, these inconsistencies should be queried and resolved by the sites.

- More than one oral anticoagulant was entered at a single time point
- A different oral anticoagulant was entered at hospital discharge/14 days than at baseline
- Both an oral anticoagulant and a non-oral/other therapy were entered at a single time point
- Different non-oral therapies were entered at hospital discharge/14 days than at baseline

6.2 IMPORTANT PROTOCOL VIOLATIONS

Patients with important protocol violations (IPVs) will be documented. The list of all protocol violations (or deviations) and their handling rules are documented in detail in the separate document Protocol Deviation Specification.

Additional IPVs may be defined during the course of the study, based on assessments at the MQRMs; if the list of IPVs needs to be enlarged, it will be documented in the report planning meeting minutes before database lock.

6.3 PATIENT SETS ANALYSED

The planned analysis for Objective 1 will be based on all eligible patients (i.e. all patients who fulfil all inclusion criteria and no exclusion criteria).

6.5 POOLING OF CENTERS

Centre will not be used as a variable in any analysis; therefore this section is not relevant.

6.6 HANDLING OF MISSING DATA AND OUTLIERS

For Objective 1, the presence of missing data is expected to be low, since the variables are expected to be collected in routine clinical practice and should therefore be available in the patient's medical record. As a result, any missing data will be treated as missing and no imputation will be performed.

The amount of missing data will be reported in the descriptive summary tables (i.e. a 'Missing' category will be displayed for variables with missing data).

Depending on the amount of missing data (e.g., >10% for a key variable), sensitivity analysis may be performed, where the distribution of other variables (age, sex, and other patient characteristics) will be compared between patients with and without missing data for each variable, and reported.

6.7 BASELINE, TIME WINDOWS AND CALCULATED VISITS

The planned analysis for Objective 1 is cross-sectional in nature. However, in addition to the data collected at the baseline visit, additional data at hospital discharge or 14 days after the baseline visit (whichever later) will be collected and included in the analysis. This is because the treatment with some NOACs (e.g. dabigatran and edoxaban) requires 5 to 10 days of pre-treatment with parenteral therapy.

7. PLANNED ANALYSIS

Summary statistics for continuous variables will include the number of observations (N), mean, standard deviation (SD), minimum, Q1 (25^{th} percentile), median, Q3 (75^{th} percentile), and maximum; tabulations of categorical variables will present all possible categories and will display the number of observations per category as well as percentages. The default precision for percentages should be one decimal place, unless otherwise specified. In general, percentages will be based on all patients in the analysis population, whether they have missing data or not. A category 'Missing' will be displayed whenever at least one patient falls within this category. For variables based on nested questions (i.e. questions which ask for more detailed information only within the subgroup of patients with a specific value for another higher lever question), percentages will be based on the patients with the corresponding value for the higher level question.

Individual patient listings will be created, which generally will be sorted by anticoagulation treatment at baseline, region, country, center number and subject number, and be based on the analysis population used for the corresponding table.

7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Analysis of demographic and other baseline characteristics is defined as a main analysis for Objective 1. Details regarding the main analysis can be found in Section 7.6.

7.2 CONCOMITANT DISEASES AND MEDICATION

Analysis of baseline concomitant diseases and medication is defined as a main analysis for Objective 1. Details regarding the main analysis can be found in Section 7.6.

7.3 TREATMENT ADHERENCE

Not applicable for Objective 1 of this study.

7.4 METHODS ADDRESSING BIAS

See study protocol.

7.5 METHODS ADDRESSING CONFOUNDING/EFFECT MEASURE MODIFICATION

Not applicable for Objective 1 of this study.

7.6 MAIN ANALYSES

Descriptive summaries of baseline patient characteristics as specified in Section 5 will be created for the overall population and by anticoagulation treatment. These analyses will be repeated by the subgroup variables defined in Section 6.4.

Descriptive summaries of anticoagulation treatment patterns, including the observed combinations of treatment at baseline and at hospital discharge/14 days as described in Table

6.1: 1, will be created for the overall population. These analyses will be repeated by the subgroup variables defined in Section 6.4.

As an exploratory analysis, baseline characteristics of the dabigatran patients will be summarized separately by the following dose groups, if the subgroups consist of at least 200 patients:

- 150 mg b.i.d.
- 110 mg b.i.d.
- 75 mg b.i.d.
- Other doses

7.8 EXPOSURE TIME

As no follow up is included and only the anticoagulation treatment at baseline is captured, exposure time will not be calculated.

7.9 SAFETY ANALYSIS

7.9.1 Adverse drug reactions

For Objective 1, patients will not be followed up after hospital discharge or 14 days after baseline. If the investigator becomes aware of an event at the baseline visit or when entering the data collected at the baseline visit that is deemed by the investigator to be related to a Boehringer Ingelheim drug, the investigator is responsible for entering this information in the Adverse Drug Reactions/Adverse Events with Fatal Outcome page of the eCRFs. However, these data will not be analyzed for Objective 1. A listing of all reported ADRs/fatal AEs will be provided in the study report appendix.

7.9.2 Laboratory data

All laboratory variables will be summarized descriptively. Creatinine clearance (mL/min) will be derived based on the collected serum creatinine data using the Cockcroft-Gault formula [3].

7.9.3 Vital signs

Weight, height, blood pressure (systolic and diastolic), heart rate and respiratory rate will be collected at baseline and summarized descriptively.

7.9.4 ECG

Electrocardiogram (ECG) assessments are not collected as part of this study.

7.9.5 Others

No other assessments or data collection are planned for Objective 1.

7.10 INTERIM ANALYSES

It is planned that one interim analysis will be performed for Objective 1 based on data from approximately the first 3,000 patients. The interim analysis will include all the relevant analyses described in this SEAP.

The final analysis will be performed once the data collection for all enrolled patients is completed, the data sets are cleaned and the database is locked for Objective 1.

8. REFERENCES

1	*CPMP/ICH/363/96:* "Statistical Principles for Clinical Trials", ICH Guideline Topic E9, Note For Guidance on Statistical Principles for Clinical Trials, current version.
2	ISPE Guidelines of Good Pharmacoepidemiology Practice, Pharmacoepidemiology and Drug Safety 2008; 17: 200–208
3	Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron. 1976;16(1):31-41.

10. HISTORY TABLE

Table 10: 1 History table

Version	Date (DD-MMM-YY)	Author	Sections changed	Brief description of change
Final	05-APR-17		None	This is the final SEAP without any modification

Statistical and Epidemiological Analysis Plan - Objective 2

c28215658-01

BI Trial No.:	1160.188
Title:	ChaRactErization of patients following aCute venous thrOmboembolism (VTE) and assessment of safety and effectiveness of dabigatran etexilate (DE) in the tReatment and secondarY prevention of acute deep vein thrombosis (DVT) and pulmonary embolism (PE) in comparison to vitamin K antagonist (VKA) in routine clinical practice - RE-COVERY DVT/PE Including protocol amendment 2 [c03118416-03]
Investigational Product(s):	Dabigatran etexilate
Responsible trial statistician(s):	Phone: Fax: Phone: Fax:
Responsible trial epidemiologist:	Phone: ; Fax:
Date of statistical and epidemiological analysis plan:	29 May 2019 SIGNED
Version:	Final

1. TABLE OF CONTENTS

LIST OF TABLES

2. LIST OF ABBREVIATIONS

Term	Definition / description
ACS	Acute coronary syndrome
AE	Adverse Event
ADR	Adverse Drug Reaction
BI	Boehringer Ingelheim
b.i.d.	Bis in die (two times a day)
BMI	Body Mass Index
DE	Dabigatran etexilate
DVT	Deep vein thrombosis
eCRF	Electronic Case Report Form
ePS	Estimated Propensity Score
ICH	International Conference on Harmonisation
IPD	Important protocol deviation
IPTW	Inverse probability of treatment weighting
ISPE	International Society for Pharmacoepidemiology
ISTH	International Society on Thrombosis and Haemostasis
ITT	Intention-to-treat
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MQRM	Medical Quality Review Meeting
NCB	Net clinical benefit
NIS	Non-interventional study
PACT-Q2	Satisfaction with anticoagulation treatment
PD	Protocol deviation
PE	Pulmonary embolism
Q1	Lower quartile
Q3	Upper quartile
SD	Standard deviation
SEAP	Statistical and epidemiological analysis plan
STEMI	ST elevation myocardial infarction
TTR	Time in therapeutic range
UA	Unstable Angina

Term	Definition / description
ACS	Acute coronary syndrome
VKA	Vitamin K antagonist
VTE	Venous thromboembolism
WHO	World Health Organization

3. INTRODUCTION

As per ICH E9 [1 and the ISPE Guidelines of Good Pharmacoepidemiology Practice [2], the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and to include detailed procedures for executing the statistical analysis of the data. This document describes the final analysis at the end of study for Objective 2.

This SEAP assumes familiarity with the non-interventional study (NIS) protocol including Protocol Amendments. In particular, the SEAP is based on the planned analysis specification as written in the NIS protocol Section 9 “Research Methods”. Therefore, SEAP readers may consult the NIS protocol for more background information on the study, e.g., on study objectives, study design and population, treatments, definition of measurements and variables, planning of sample size.

SAS® Version 9.3 or higher will be used for all analyses.

4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

None.

5. OUTCOMES

Objective 1 of the RE-COVER was described in a separate Document, "Statistical and Epidemiological Analysis Plan - Objective 1".

The following two questions were investigated in the SEAP for Objective 1:

- With the approval of dabigatran etexilate and other non-vitamin K anticoagulants (NOACs), what treatments are being administered for acute VTE in routine clinical practice in the different regions of the world?
- When presenting with an acute VTE event, what factors influence the choice of treatment for the event?

Objective 2 of the RE-COVER study is designed to analyze the safety and effectiveness of dabigatran etexilate regimens in the treatment of DVT and PE over 1 year of follow-up in comparison to a VKA regimen.

Additional further exploratory objectives are:

- To evaluate geographical variations in health care resource utilization and patient satisfaction with their treatment in routine clinical practice.
- To conduct stratified analysis for the safety and effectiveness of dabigatran etexilate according to dose.

The following three questions will be investigated:

- What is the safety and effectiveness of dabigatran etexilate versus VKA under conditions of routine clinical practice?
- What health care resources are utilized in patients who are treated with dabigatran etexilate or VKA for an acute VTE under conditions of routine clinical practice?
- For patients who are treated with either dabigatran etexilate or VKA, what is patients' satisfaction with their treatment under conditions of routine clinical practice?

5.1 MAIN OUTCOMES

The main outcomes for Objective 2 are:

- Primary safety outcome measure: ISTH Major bleeding and ISTH clinically relevant non major bleeding (CRNMB)
- Primary effectiveness outcome measure: symptomatic recurrent VTE including VTE-related mortality

A patient is considered to have major bleeding if the bleeding is fatal and/or meets at least one of the following criteria:

- Overt bleeding associated with a reduction in hemoglobin of at least 20 grams per liter or leading to a transfusion of at least 2 units of blood or packed cells

- Symptomatic bleeding in a critical area or organ: intraocular, intracranial, intraspinal, or intramuscular with compartmental syndrome, retroperitoneal bleeding, intraarticular bleeding or pericardial bleeding

A patient is considered to have clinically relevant non major bleeding if the bleeding caused at least one of the following actions:

- Leading to hospitalization or increased the level of health care
- Requiring medical intervention by a healthcare professional
- Prompting a face to face evaluation

5.2 SECONDARY OUTCOMES

The following secondary outcomes will be analysed from the time of diagnosis of the index event to 12 months post diagnosis:

- Recurrent DVT and/or PE
- VTE-related mortality
- All-cause mortality

6. GENERAL ANALYSIS DEFINITIONS

6.1 EXPOSURE

Patients will be analyzed in treatment groups (DE or VKA) according to the treatment that has been prescribed at the baseline visit and at hospital discharge or 14 days after the baseline visit and recorded in the Patient Status eCRF (obtained at enrolment). The start of the follow up will be defined as the initiation of each of these two treatments.

6.2 IMPORTANT PROTOCOL DEVIATIONS

Patients with important protocol deviations (IPDs) will be documented. The list of all protocol deviations (PDs) and their handling rules are documented in detail in the separate document Protocol Deviation Specification. Additional IPDs may be defined during the course of the study, based on assessments at the MQRMs; if the list of IPDs needs to be enlarged, it will be documented in the report planning meeting minutes before database lock.

6.3 PATIENT SETS ANALYSED

Three patient sets are defined: "all eligible", “all eligible treated”, and “restricted”; the following rules are used to assess if a patient is included in an analysis set:

ALL ELIGIBLE: If a patient fulfills all inclusion criteria and no exclusion criteria.

ALL ELIGIBLE TREATED: If a patient fulfills all inclusion criteria and no exclusion criteria and took prescribed treatment at least once.

RESTRICTED: The restricted population is composed of all eligible patients within the trimmed patient set which lie within the region of propensity score overlap (for definition see Section 7). As the propensity score analysis will be based on eligible patients who at least once took the prescribed treatment, the restricted population excludes patients prescribed the treatment but who did not take it.

Number of patients included in the different patient sets will be shown by treatment group.

6.5 POOLING OF CENTRES

Centre will not be used as a variable in any analysis.

6.6 HANDLING OF MISSING DATA AND OUTLIERS

Every reasonable attempt will be undertaken to ensure completeness of data collection.

A review of the data will be performed at the MQRM in order to review reasons and patterns of missing values in the data and to decide on the need for countermeasures.

In general, "unknown" data will be treated as missing. This is because when in the CRF the option "unknown" is given for a question, it is obligatory to tick one optional answer (no missing values). If no "unknown" option is present in the respective question, it is not obligatory to tick one optional answer (leading to missing values if not answered at all).

In the descriptive analyses, unknown values will be treated as an additional category.

Multiple imputations (MI) technique will be used for handling missing value for any of the covariates included in the logistic regression (i.e., alcohol abuse, BMI, Creatinine clearance and smoking status) for the derivation of the average estimated propensity score (ePS). The mean of the ePS for each patient will be estimated from each of the m imputed data sets based generally on Rubin's approach for combining multiple imputed estimates. For example, if 5 imputed data sets are created each patient will have 5 ePS from which a mean ePS would be directly estimated for each patient.

6.7 BASELINE, TIME WINDOWS AND CALCULATED VISITS

The baseline analysis for Objective 2 is cross-sectional in nature. Patients will be assigned to the two treatment groups as described in Section 6.1. For the descriptive analysis, in addition to the data collected at the baseline visit, additional data at hospital discharge or 14 days after the baseline visit (whichever is later) will be collected and summarized descriptively. For the logistic regression analysis to produce estimated propensity scores, data up to the start of dabigatran or VKA treatment will be considered as baseline and will be used as model covariates.

Data from the longitudinal follow-up for collected laboratory data, healthcare resource utilization, and PACT-Q2 will be summarized descriptively at the time points given in Table 5.3: 2. In case time windows will be needed in order to assign observations to visits, observations from 45 days to 134 days after diagnosis will be assigned to month 3, from 135 days to 269 days after diagnosis will be assigned to month 6 and more than 269 days will be assigned to year 1. In case more than one observation is assigned to a certain visit the observation closest to the calculated visit will be assigned. In case two or more observations are on the same day the last will be assigned. Month 3 corresponds to 91 days, month 6 to 182 days, and year 1 to 365 days.

7. PLANNED ANALYSIS

The propensity scores will be estimated after the multiple imputation is performed based on eligible patients who took the prescribed treatment at least once. A logistic regression model will be used to model the probability of receiving dabigatran etexilate in comparison to vitamin K antagonist, i.e. high propensity scores, close to 1, indicate a higher probability of dabigatran use. See Section 7.6 for details.

For the main analysis, only those patients will be included who have a shared range of ePS. This patient set is denoted as the **restricted patient set in the SEAP**. Patients with ePS outside the overlapped ePS range will be trimmed (or excluded).

Safety analyses will be conducted for the all eligible treated patient set.

Summary statistics for continuous variables will include the number of observations (N), mean, standard deviation (SD), minimum, Q1 (25th percentile), median, Q3 (75th percentile), and maximum; tabulations of categorical variables will present all possible categories and will display the number of observations per category as well as percentages. The default precision for percentages should be one decimal place, unless otherwise specified. In general, percentages will be based on the number of patients eligible for the particular analysis, whether they have missing data or not. A category 'Missing' will be displayed whenever at least one patient falls within this category.

Data collected in the baseline eCRFs (informed consent, patient status, demographics, inclusion/exclusion criteria, vital signs and physical assessment, clinical signs and symptoms) will be considered as respective baseline values. Data collected in the hospital discharge related eCRFs will be summarized for the hospital discharge time point. Despite the fact that the analysis is of cross-sectional nature, data collected in eCRFs at hospital discharge will also be summarized in the descriptive analysis.

7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Analysis of demographic and other baseline characteristics (Table 5.3: 1) is defined as a further analysis for Objective 2. Details regarding the further analysis can be found in Section 7.7.

7.2 CONCOMITANT DISEASES AND MEDICATION

Analysis of baseline concomitant diseases, concomitant diseases which started before VTE diagnosis, medical history, and concomitant medication will also be conducted. Concomitant diseases and medical history will be coded using the latest version of MedDRA dictionary, and medications will be coded using the latest WHO drug dictionary preferred name. Both will be summarized using frequencies and percentages for all eligible patients. Patients taking more than one medication in the same preferred name will be counted once.

7.3 TREATMENT ADHERENCE AND PERSISTENCE

Treatment persistence will be assessed as time from treatment onset until permanent treatment discontinuation and summarized descriptively.

Time in therapeutic range (TTR) [3] calculated from INR results obtained at the respective timepoints will be analyzed. For each patient who took prescribed VKA, the TTR percentage will be calculated by linear interpolation between successive test results, and then the number of days in the therapeutic range will be estimated. For the target range 2.0 to 3.0 will be considered.

7.4 METHODS ADRESSING BIAS

See study protocol.

7.5 METHODS ADRESSING CONFOUNDING/EFFECT MEASURE MODIFICATION

See Section 7.6.

7.6 MAIN ANALYSIS

Analyses for the main outcome events are mainly defined for patients in the restricted patient set and, in some occasions for patients in the eligible patient set who took at least once the treatment. In general, recurrent events will be disregarded for the analyses of the outcomes.

Main outcome events – Incidence rates

For the treatment groups incidence rates with 95% CIs, cumulative risks, and KM curves will be calculated on the eligible patient set. Only patients from the eligible patient set that took the prescribed treatment at least once are considered. Patients who complete the planned anticoagulation treatment prior to one year or discontinue initial anticoagulation treatment permanently will be censored at date of last drug intake + 6 days for both DE and VKA or at first intake of another relevant anticoagulation treatment, whichever occurs first. A patient is considered to have permanently stopped initial anticoagulation treatment if other relevant anticoagulation treatment is initiated or if treatment has been interrupted for more than 30 days.

The main analysis of overall incidence rates with 95% confidence intervals, cumulative risks and KM curves will be repeated for the subgroups defined in Section 6.4. A graphical (forest-plot-like) display will additionally be prepared to summarize the subgroup results per outcome on the eligible patient set. Additional subgroup analyses might be conducted as needed.

For all important main and secondary outcomes for patients initiating VKA at baseline, and patients initiating DE at baseline, incidence rates, cumulative risks with 95%-confidence intervals will be calculated by treatment, and for a graphical comparison Kaplan Meier curves

will be plotted. These statistics will be calculated on the restricted patient set, and patients who never took the prescribed treatment will not be included. For the restricted patient set, as a sensitivity analysis, the calculation of incidence rates and cumulative risks will be performed without censoring at permanent discontinuation (intention to treat (ITT) approach). KM curves will be plotted.

Secondary and further outcome events

For the further outcome events such as Recurrent DVT and/or PE, VTE-related mortality, All-cause mortality, Life-threatening bleeding events, Acute coronary syndrome, and Net clinical benefit, incidence rates, KM estimates and KM curves will be calculated for the eligible patients who took the prescribed treatment at least once for both treatment groups. Additionally, above analyses will be performed on the restricted patient set.

Technical specifications

Incidence rates per 100 patient years with 95%-confidence intervals for important outcome events (single and composite outcome) will be calculated the following way. In addition, these analyses will be done stratified by ongoing chronic antiplatelet treatment at baseline. Recurrent events will be disregarded for calculation of incidence rates. Incidence rate λ is defined as

$$\text{Incidence rate } \lambda\left[\frac{1}{100}\ \text{pt.yrs}\right]: \\ = 100 * \frac{\text{number of patients with event}}{\frac{1}{365.25} * \sum_{\text{patients}} \text{person time at risk(event)[days]}}$$

where the denominator is the patient years under risk. The 2-sided 95% CI of λ will be calculated based on the Poisson distribution and its relationship with the Chi-square distribution according to the following formula [7]

$$[\lambda_{\text{lower}}, \lambda_{\text{upper}}] \ = \ [\ \{0.5\,\chi^2(\alpha/2;\ 2x)\}/n\ , \{0.5\,\chi^2(1-\alpha/2;\ 2x+2)\}/n\ ],$$

where $\chi^2(\gamma, \boldsymbol{r})$ is the 100λth the percentile of a chi-square distribution with $\boldsymbol{r}$ degrees of freedom, n is the patient years under risk for the event, x is the number of patients with event, and α for $(1-\alpha)$% 2-sided CIs.

Cumulative risk is defined as cumulative proportion of patients with event and will be calculated as the Kaplan-Meier estimate for failure. The Kaplan-Meier curve will be plotted. The proportion of patients with event, as well as the proportion of censored patients falling into specified intervals and the Kaplan-Meier estimate for failure will be computed for the following categories:

- <= 3 months
- 3< and <=6 months
- 6< and <=12 months
- > 12 months

Distribution and overlap of the ePS between the two groups

The ePS is the probability of a patient being in one of the two treatment groups, conditional on relevant observed baseline variables. For this analysis, the ePS for being in the DE group vs. VKA group will be estimated using the multivariable logistic regression model.

Baseline variables will be included in the model based on their known/suspected association with the main outcome events. As suggested by the literature, only confounders or variables associated with the outcomes will be considered. Table 7.6: 1 summarizes the identified variables to be included in the logistic regression model based on these criteria. In the event of non-convergence, investigations will be done to explore the most likely reason and to aggregate or exclude the responsible variables. Generally, all missing values of variables which will be included in the logistic regression model will be imputed according to Section 6.6.

Table 7.6: 1 Variables for PS model

Variable	**Comment**	**Potential for missing values and implications**
Age group	<65, 65 to <75, >=75	Very unlikely
Sex	Male, Female	Rare
BMI	<18.5, 18.5 to <25, 25 to <30, 30 to <35, ≥35, missing	Missing will be imputed
Creatinine clearance	<30, 30 to < 50, 50 to <80, ≥80, missing	Missing will be imputed
Region	Europe, Middle East, North America, Latin America, Asia, Other	Very unlikely
Smoking status	non-smoker, current smoker, past smoker, unknown	Rare – will impute missing and unknown values
Alcohol use	Never, < 1 drink/week, 1-7 drinks/week, >=8 drinks/week, unknown	Rare – will impute missing and unknown values
Trauma or surgery	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed – 3 month time window)	N/A
Immobilization	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (past and concomitant))	N/A
Obesity	True or False (CRF categories, medical history and concomitant diseases will be	N/A

	manually reviewed (past and concomitant))	
Thrombophilia	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (past and concomitant))	N/A
Pregnancy and peri-partum period	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (3 month time window))	N/A
Concomitant use of aspirin	True or False (concomitant medication and other VTE therapies will be manually reviewed (concomitant))	N/A
Concomitant use of clopidogrel	True or False (concomitant medication and other VTE therapies will be manually reviewed (concomitant))	N/A
Estrogen therapy	True or False (concomitant medication and other VTE therapies will be manually reviewed (3 month time window))	N/A
Concomitant cancer	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
Antiphospholipid antibody syndrome	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
Polycythemia	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
Hypertension	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
History of previous VTE	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (regardless of timing))	N/A
Diabetes mellitus	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
Heart failure	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (regardless of	N/A

	timing))	
History of major bleeding	True or False (medical history and concomitant diseases will be manually reviewed (regardless of timing))	N/A
Stroke	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (regardless of timing))	N/A
Chronic hepatic disease	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
Chronic renal disease	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
Index event	DVT, PE, DVT and PE (CRF)	Very unlikely
Prior gastrointestinal bleed	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (regardless of timing))	N/A
Venous insufficiency or varicose veins	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A

Overlap of estimated propensity scores

Overlap is the degree to which the ePS of the two treatment groups have a shared range. It is typically best examined graphically but will also be quantified by treatment group using the Kolmogorov Smirnov procedures [4] as a key quantitative indicator of the degree of overlap. To depict overlap in this context, a mirrored bar chart and overlapping line graphs of the estimated ePS of the two groups will be employed. One or both group distributions may be depicted before and after trimming. The same figure will be produced for logit transformed ePS.

In the following if not stated otherwise for all analyses only patients from the restricted patient set are considered.

1.) Main analysis of the main outcomes (Major bleeding and clinically relevant non major bleeding, and symptomatic recurrent VTE including VTE-related mortality)

The main analysis will be a PS analysis using overlap weights. A Cox regression model including the overlap weights will be performed. In addition, the main outcomes will be analyzed within the subgroups defined in Section 6.4, when the number of patients per treatment group is at least 500. For the overlap weights analysis each patient in the treatment groups DE and VKA will be weighted proportional to the probability of that

patient being assigned to the opposite treatment group. The weights are called "overlap weights" and are calculated as follows:

$$O_{DE} = 1 - P$$
$$O_{VKA} = P$$

where P is the estimated propensity score of a patient being assigned to the DE treatment group.

To assess baseline imbalance for the overlap weights by ePS, the mean of each covariate of interest will be calculated for patients in each of the treatment groups (DE and VKA). For binary variables, the proportion will be used instead. For categorical variables with more than two categories, dummy binary variables will be derived to represent each category and the proportion calculated for each of the treatments.

The standardized differences will be calculated for the mean for each covariate and the standardized differences (denoted as d) for the mean of a continuous variable is defined [10] as:

$$d = \frac{(\bar{x}_{DE} - \bar{x}_{VKA})}{\sqrt{\frac{s^2_{DE} + s^2_{VKA}}{2}}}.$$

where $\bar{x}_{DE}$ and $\bar{x}_{VKA}$ denote the sample mean of the covariate in DE and VKA patients, respectively, while s^2_{DE} and s^2_{VKA} denote the sample variance of the covariate in DE and VKA patients, respectively.

For dichotomous variables, the standardized difference is defined as:

$$d = \frac{(\hat{p}_{DE} - \hat{p}_{VKA})}{\sqrt{\frac{\hat{p}_{DE} * (1 - \hat{p}_{DE}) + \hat{p}_{VKA} * (1 - \hat{p}_{VKA})}{2}}}$$

where $\hat{p}_{DE}$ and $\hat{p}_{VKA}$ denote the proportion or mean of the dichotomous variables in DE and VKA patients, respectively. This approach described above will be repeated for the same covariates of interest after applying the appropriate overlap weight to each patient based on the ePS and will be presented for each covariate. Note that after weighting, the mean will be used instead of the proportion for binary and categorical variables because the weighting applied will make the variables continuous. After weighs will have been applied sample means and variances will be substituted by weighted means and weighted sample variances denoted by

$$\bar{x}_{weight} = \frac{\sum O_i x_i}{\sum O_i}$$

and

$$s^2_{weight} = \frac{\sum O_i}{(\sum O_i)^2 - \sum O_i^2} * \sum O_i (x_i - \bar{x}_{weight})^2$$

where O_i is the weight assigned to the i-th patient.

These unadjusted (i.e., prior to using the overlap weights with ePS adjustment) and adjusted (using overlap weights) standardized differences for the mean will be presented for each covariate.

Balance will then be assessed to determine whether the weighted covariates procedure produces standardized differences for the mean and SD closer to zero (indicating a better balance between the study cohorts) than those of unadjusted covariates. Adjusted standardized differences for any covariate that are greater than 0.25 will potentially warrant re-estimation of the ePS model to improve balance.

2.) Sensitivity analysis of the main outcomes (Major bleeding and clinically relevant non major bleeding, and symptomatic recurrent VTE including VTE-related mortality)

a.) As sensitivity analysis for the comparison of the two treatment groups, dabigatran etexilate and VKA, with regards to the main outcomes a multivariable Cox regression model will be performed. The multivariable Cox regression model will include the variables given in Table 7.6: 2. In case of model convergence issues due to limited number of events, some variables may be removed from the model as deemed necessary.

Table 7.6: 2 Variables for Cox regression model

Variable	Comment	Potential for missing values and implications
Age	continuous	Very unlikely
BMI	continuous	Missing will be imputed
Trauma or surgery	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed – 3 month time window)	N/A
Immobilization	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (past and concomitant))	N/A
Thrombophilia	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (past and concomitant))	N/A
Pregnancy and peri-partum period	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (3 month time window))	N/A

Concomitant use of aspirin	True or False (concomitant medication and other VTE therapies will be manually reviewed (concomitant))	N/A
Concomitant use of clopidogrel	True or False (concomitant medication and other VTE therapies will be manually reviewed (concomitant))	N/A
Estrogen therapy	True or False (concomitant medication and other VTE therapies will be manually reviewed (3 month time window))	N/A
Concomitant cancer	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
Hypertension	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A
History of previous VTE	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (regardless of timing))	N/A
History of major bleeding or prior gastrointestinal bleed	True or False (medical history and concomitant diseases will be manually reviewed (regardless of timing))	N/A
Chronic renal disease	True or False (CRF categories, medical history and concomitant diseases will be manually reviewed (concomitant))	N/A

Patients who complete the planned anticoagulation treatment prior to one year or discontinue initial anticoagulation treatment permanently will be censored at date of last drug intake + 6 days for both dabigatran etexilate and VKA or at first intake of another relevant anticoagulation treatment, whichever occurs first. A patient is considered to have permanently stopped initial anticoagulation treatment if other relevant anticoagulation treatment is initiated or if treatment has been interrupted for more than 30 days.

Multiple imputation methods (e.g. the fully conditional specification method) will be used to impute both continuous and categorical variables appropriately with arbitrary missing patterns.

b.) In addition, Cox regression analysis (based on the main analysis model) will be performed without censoring at permanent discontinuation (intention to treat (ITT) approach).

3.) Analyses of secondary and further outcome events

The same Cox regression model as used for the primary outcomes will be used for the secondary and further outcomes, if enough events are reported. The ePS overlap weights analysis on the specific outcome event might be performed for all outcomes.

7.8 EXPOSURE TIME

Exposure data will be summarized based on the eligible and restricted patient sets for dabigatran etexilate and VKA. Prescribed but never taken patients will be excluded from all the analyses described in this section.

The total exposure time will be calculated for each treatment group X as:

Exposure treatment X[days]=last date of treatment intake – first date of treatment intake + 1

The calculated days are converted to months (12*days/365.25) for summary tables.

In addition, total patient years will be calculated and presented:

Total patient − years treatment X[years] = $\sum_{\text{Patients}}$ Exposure treatment X [days]/365.25.

7.9 SAFETY ANALYSES

7. 9.1 Adverse events

Analyses for AEs will be based on the all eligible treated patient set. The analysis of adverse events will be descriptive in nature and will be based on BI standards (see DM&SM: Handling and summarization of adverse event data for clinical trial reports and integrated summaries [6]). Adverse events will be coded with the most recent version of MedDRA. The summary of standard AE tables will be based on the AE/ SAE CRFs. According to the reporting procedures for this study, all SAEs and outcome events are to be documented on the SAE or outcome events form irrespective of causal relationship to antithrombotic treatment or any other BI drug. Non-serious AEs which are not deemed related to antithrombotic treatment or to any other BI drug should not be included in the CRF (see study protocol, Section 11.2).

For the analysis, only AEs which are considered treatment-emergent for antithrombotic treatment are included.

For analysis of adverse drug reactions (ADRs), all AEs which are related to antithrombotic therapy according to the investigators' judgment are to be included.

The following tables will be created by treatment group:

- Overall summary of ADRs.
- Frequency [N (%)] and incidence rate (per 100 patient years) of patients with ADRs (related to antithrombotic therapy).
- Frequency [N (%)] and incidence rate (per 100 patient years) of patients with ADRs (related to antithrombotic therapy) leading to discontinuation of antithrombotic treatment.
- Frequency [N (%)] and incidence rate (per 100 patient years) of patients with ADRs (related to antithrombotic therapy) leading to death.
- Frequency [N (%)] and incidence rate (per 100 patient years) of patients with Serious ADRs (related to antithrombotic therapy).
- Frequency [N (%)] and incidence rate (per 100 patient years) of patients with fatal AEs

7. 9.2 Laboratory data

The following laboratory variables at the described time points (Table 5.3:) will be summarized descriptively.

Table 7.9.2: 1 Laboratory Variables

Parameter	**Unit**
Serum Creatinine	mg/dL
Hemoglobin	g/L
International Normalized Ratio (INR)	

Creatinine Clearance	mL/min

Creatinine clearance (mL/min) will be derived based on the collected serum creatinine data using the Cockcroft-Gault formula [55].

7. 9.3 Vital signs

Height, blood pressure (systolic and diastolic), heart rate and respiratory rate will be collected at baseline and summarized descriptively. Weight at the described time points in Table 5.3: will be summarized descriptively.

7. 9.4 ECG

Electrocardiogram (ECG) assessments are not collected as part of this study.

7. 9.5 Others

None.

8. REFERENCES

1	*CPMP/ICH/363/96:* "Statistical Principles for Clinical Trials", ICH Guideline Topic E9, Note For Guidance on Statistical Principles for Clinical Trials, current version.
2	ISPE Guidelines of Good Pharmacoepidemiology Practice, Pharmacoepidemiology and Drug Safety 2008; 17: 200–208
3	Rosendaal FR, Cannegieter SC, van der Meer FJ, Briet E. A method to determine the optimal intensity of oral anticoagulant therapy. Thromb Haemost 1993; 69: 236–39.
4	An Investigation of the Kolmogorov-Smirnov Two Sample Test using SAS®. Cardinal Health, Advanced Analytics Team, Dublin, Ohio 43017, USA
5	Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron. 1976;16(1):31-41.
6	001-MCG-156: "DM&SM: Handling and summarization of adverse event data for clinical trial reports and integrated summaries"); current version, IDEA for CON.
7	Hahn GJ, Meeker WQ. Statistical Intervals: A Guide for Practitioners. Wiley, New York/Chichester/Brisbane/Toronto/Singapore: John Willey and Sons; 1991.
8	Til Stürmer,* Kenneth J.Rothman, Jerry Avorn, and Robert J. Glynn. Treatment Effects in the Presence of Unmeasured Confounding: Dealing With Observations in the Tails of the Propensity Score Distribution- A Simulation Study. Am J Epidemiol. Oct 1, 2010; 172(7): 843-854
9	Li, Fan & Lock Morgan, Kari & Zaslavsky, Alan. (2014). Balancing Covariates via Propensity Score Weighting.
10	Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies, Statist. Med. 2015, 34 3661–3679

10. HISTORY TABLE

Table 10: 1 History table

Version	Date (DD-MMM-YY)	Author	Sections changed	Brief description of change
Final	**29-May-19**		None	This is the final SEAP without any modification