Comparison of **O**ral anticoagulants for extended **Ve**nous **T**hromboembolism

(COVET)

Statistical Analysis Plan
NCT 03196349
July 11, 2018

Prepared by Anna Giczewska

Duke Clinical Research Institute

Table of contents

1. Background

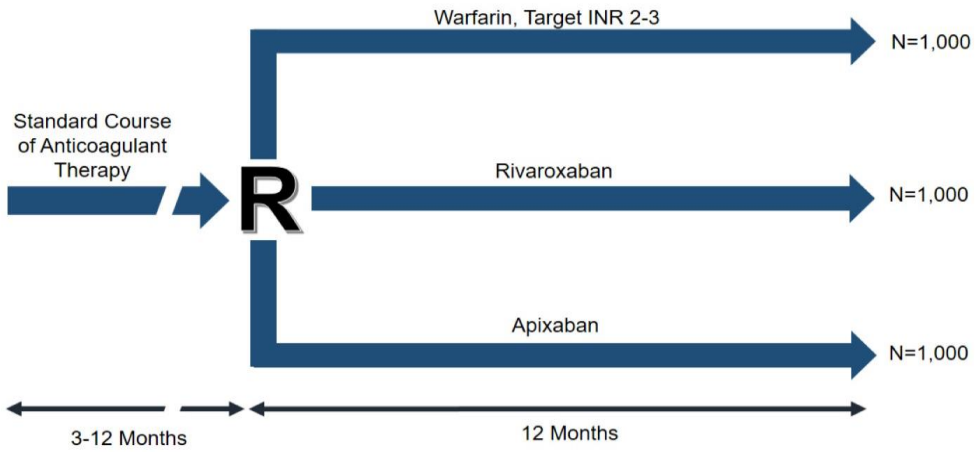
Venous thromboembolism (VTE), including deep vein thrombosis (DVT) and pulmonary embolism (PE), is a common, potentially fatal, yet treatable condition and is the third leading cause of mortality by cardiovascular disease. The standard therapy for patients with VTE is anticoagulation for 3-6 months; the most common oral anticoagulants used are vitamin K antagonists (VKA) and direct oral anticoagulants (DOACs). After completing 3-12 months of anticoagulation, patients, clinicians and policy makers are faced with the crucial decision of whether or not to continue anticoagulation lifelong and which anticoagulant to choose. Vitamin K antagonists were the only oral anticoagulants available for many years and DOACs have expanded the choices now available.

To compare a DOACs (apixaban and rivaroxaban) to VKA (warfarin) for extended anticoagulation for safety and efficacy, we propose a pragmatic clinical trial in which 3000 patients who have confirmed acute symptomatic and unprovoked DVT/PE, have completed initial treatment of oral anticoagulation therapy for 3-12 months and are at high risk for recurrent VTE will be randomly assigned in a 1:1:1 ratio to warfarin, apixaban or rivaroxaban. Maximum follow-up will be 12 months. The primary efficacy endpoint is rate of recurrent VTE. The primary safety endpoint is clinically relevant bleeding and/or clinically relevant non major bleeding.

Further details on the background and significance of this trial can be found in the clinical trial study protocol.

2. Study Design

The COVET study is a multi-center, randomized, open-label study conducted in the United States and Canada of approximately 3000 patients who have confirmed acute symptomatic and unprovoked DVT/PE, have completed initial treatment of oral anticoagulation therapy for 3-12 months and are at high risk for recurrent VTE. Eligible patients who consent to participate in the study will be randomized to warfarin, apixaban, or rivaroxaban (see Figure).



3. Study Objectives

The DCRI statistics group will investigate the following objectives:

Primary Safety Objectives:

- Determine if apixaban is superior to warfarin in the reduction of clinically relevant bleeding.
- Determine if rivaroxaban is superior to warfarin in the reduction of clinically relevant bleeding

Primary Efficacy Objectives:

- Determine if apixaban is non-inferior to warfarin in the prevention of recurrent venous thromboembolism.
- Determine if rivaroxaban is non-inferior to warfarin in the prevention of recurrent venous thromboembolism.

Exploratory Objective:

- An exploratory descriptive comparison of apixaban versus rivaroxaban for the prevention of clinically relevant bleeding and recurrent VTEs as a secondary objective.

4. Analysis Population

All-randomized (Intention To Treat - ITT) – this population includes all randomized participants regardless of the actual treatment received. This population will be used for the primary analysis of the primary safety and efficacy endpoints and a secondary analysis of the primary safety endpoints.

Per protocol population – this population excludes participants who have not started the assigned therapy by the time of the 1 month telephone interview. For time-to-event analyses, participants who stop taking the assigned treatment will be censored after permanently stopping the assigned treatment. This population will be used for the primary analysis of the primary safety endpoints and a secondary analysis of the primary efficacy endpoints.

5. Data Sources

There will be only one data source eClinical Operating System (eCOS), which will include:

- electronic Case Report Form (eCRF) which will complete by sites representatives and DCRI Call Center
- CEC database which will be completed by CEC reviewers.

6. Endpoints

6.1 Primary Endpoints

6.1.1 Primary Safety Endpoints

Clinically relevant bleeding (composite of major bleeding (MB) and/or clinically relevant non major bleeding (CRNMB)) is the primary safety outcome which specifically addresses

concerns relevant to all patients with VTE. For health care providers and stakeholders, this outcome is relevant as it has been the source of recent law suits in the U.S. concerning DOAC-associated bleeding. Our criteria to define clinically relevant bleeding are consistent with those in the aforementioned VTE trials and published by the International Society on Thrombosis and Haemostasis (ISTH) (Kaatz 2015).

Primary safety objectives in patients at high-risk for recurrent VTE, the COVET trial will:

- Determine if apixaban is superior to warfarin, target INR 2-3, in the reduction of clinically relevant bleeding.
- Determine if rivaroxaban is superior to warfarin, target INR 2-3 in the reduction of clinically relevant bleeding.

6.1.2 Primary Efficacy Endpoint

Recurrent VTE is the primary efficacy outcome, defined according to criteria described in the 2012 Chest Guidelines for Diagnosis of DVT (Bates 2012) and 2014 European Society of Cardiology Guidelines for Diagnosis of PE (Konstaninides 2014). This outcome is patient centered and important for all stakeholders given the clinical (symptoms of pain and discomfort), social, and economic burden associated with it.

Primary sefficacy objectives in patients at high-risk for recurrent VTE, the COVET trial will:

- Determine if apixaban is non-inferior to warfarin, target INR 2-3, in the prevention of recurrent venous thromboembolism.
- Determine if rivaroxaban is non-inferior to warfarin, target INR 2-3, in the prevention of recurrent venous thromboembolism.

6.1.3 Other Study Endpoints

- All-cause mortality
- Major bleeding
- Clinically relevant non-major bleeding
- Premature termination of study medication
- Vascular events (myocardial infarction, ischemic stroke)

These study endpoints will be defined in detail in the Clinical Event Classification (CEC) charter.

6.2 Secondary Objectives

In patients at high-risk for recurrent VTE, the COVET trial will

- Compare the rates of clinically relevant bleeding between apixaban and rivaroxaban.

7. Sample Size Justification

The planning and sample size calculations for the COVET study were designed to compare warfarin vs. apixaban and warfarin vs. rivaroxaban. A direct exploratory and descriptive comparison of the two DOACs will be conducted but the study was not sized to be adequately powered to detect statistically significant differences between those treatment groups.

For the primary safety endpoint, there will be two comparisons, apixaban versus warfarin and rivaroxaban versus warfarin, each of which will be tested at the two-sided 0.05 level of significance to detect a 50% reduction (HR=0.50) in clinically relevant bleeding compared to warfarin. For the primary efficacy endpoint, there will be two comparisons, apixaban versus warfarin and rivaroxaban versus warfarin to determine is the DOACs are non-inferior to warfarin in the prevention of recurrent venous thromboembolism. The non-inferiority assessments will be based on comparing the one-sided 95% upper confidence interval with a 2% increase in recurrent VTE. The type I error rates for the non-inferiority and superiority analyses have been set at 0.05 one-sided and 0.05 two-sided, respectively. These are considered standard values for two arm studies. We have elected to not alter them for this three-arm design because the clinical questions of interest are the two comparisons of the DOACs versus warfarin.

For the primary efficacy analysis, using an outcome of recurrent VTE, we expect 12-month rate of 0.87% for the warfarin arm compared with 1.5% for the DOACs (Einstein Investigators 2010; Agnelli 2013). Sample sizes of 950 per arm are sufficient to provide 80% power to have the upper 95% one-sided confidence interval less than 2%. The proposed sample size of 1000 participants per arm allows for 5% missing data due to loss-to-follow-up, deaths not related to VTE, and lack of starting the assigned study drug. These calculations were obtained using a simulation method with nQuery Advisor 7.0 software.

For the primary safety analysis, the RE-MEDY study reported clinically relevant bleeding rates of 10.2% for the warfarin treated patients (Schulman 2013). Using a 12-month rate of 6.8%, a per group sample size of 950 participants (or 1000 / group with an allowance for 5% missing data) will provide greater than 90% power to detect a 50% event rate reduction for the DOAC arm. Similarly, assuming a lower 12-month event rate of 6.0% for the warfarin arm, the sample size of 950 participants per group will provide 88% power to detect a 50% event rate reduction for the DOAC arm. These calculations assume a two-sided 0.05 type I error rate and are based on a Cox model.

8. Data Safety Monitoring Board (DSMB)

The Data Safety and Monitoring Board (DSMB) will convene approximately every 6 months to evaluate the progress of the trial and review accumulating safety data. Details on the roles and responsibilities of the DSMB are outlined in the DSMB charter. Summary reports and recommendations from the meetings will be distributed to the PIs within a reasonable period of time following the meeting. There are no formal stopping rules for safety in the trial.

8.1 Interim Analysis

There is no interim analysis planned for this trial.

9. Statistical Methodology

9.1 General Analysis Conventions

9.1.1 Statistical significance

Statistical comparisons will be performed using two-sided significance tests. The level of significance for all secondary efficacy and safety analyses will be set at an alpha level of 0.05.

9.1.2 Assumption checks

Before final analysis, the basic assumptions underlying the planned approach, including the validity of the proportional hazards assumption for the Cox model, will be checked and transformations or nonparametric tests will be used as needed.

9.1.3 Descriptive statistics

Mean, standard deviation, median, 25^{th} percentile, 75^{th} percentile, minimum, and maximum will be presented for continuous variables. Categorical variables will be presented as a simple count and percentage. Other descriptive statistics may be provided where appropriate. If necessary, plots of descriptive statistics will also be provided.

9.1.4 Study Tables and Listings

Most study data will be presented by treatment group and overall. Data may also be listed by subject id and by site where appropriate.

9.1.5 Validation procedures

All programs to create analysis datasets and perform analyses will be generated and validated under the guidance of the DCRI Statistical SOPs.

9.1.6 Subgroup analyses

The study investigators have identified several pre-specified subgroups of interest. Those subgroups include: initial anticoagulant, race, sex, renal impairment, obesity, and antiplatelet use (Buckner 2012; Liao 2014; White 2009). Given the expected low number of events for both primary endpoints, the subgroups will need to be interpreted cautiously. For each subgroup we will compute an interaction p-value and then estimate hazard ratios for each subgroup.

9.1.7 Timing of analyses

The need for more than one database lock is not anticipated for this trial. The analyses of the safety and efficacy data will be conducted within a reasonable time following the database lock.

9.1.8 Visit Windows

Subjects should be seen for baseline evaluation and randomization. All visits except baseline visit will be phone calls.

Time (Day)	Screening / Randomization	Month 1*	Month 6*	Month 12*
Visit	Visit 1	Visit 2	Visit 3	Visit 4
Visit Window	30 days	(+/- 1 week)**	(+/- 4 week)***	(+ 8weeks)
Type of Visit	In-person	Phone call	Phone call	Phone call

*Canadian participants will be contacted by the Canadian enrolling centers to complete phone assessments.

**If the participant is not reached within the +/- 1 week window, please attempt contact until Month 3.

***If the participant is not reached within the +/- 4 week window, please attempt contact until Month 9.

9.2 Analysis of Baseline Data

Baseline demographic and clinical characteristic descriptive data will be summarized in accordance with the general analysis conventions. If treatment group comparisons are warranted, differences will be examined using a two-sample Student t-test for continuous variables (or Wilcoxon rank sum test if the assumptions are not met) and chi-square test for categorical variables (or Fisher's exact test).

The following baseline data will be included in the final report:

- Demographic characteristics (e.g. sex, ethnicity, race, and age)
- Medical History characteristics (e.g. unprovoked deep vein thrombosis, unprovoked pulmonary embolism, GI bleed)
- Prior medications (e.g. Ketoconazole, Iitraconazole, Aspirin, Clopidogrel)

9.3 Analysis of Time to "Response" Data

9.3.1 Primary and Secondary Endpoint Analysis

The primary analyses for safety primary endpoint will be based on a stratified Cox model. The ascertainment time will be based on the time of the first event or censoring time. Stratification will be based on the anticoagulant used during the screening period. The all-randomized analysis dataset will be used for the primary analyses. The treatment effect for each DOAC compared to warfarin will be computed using the estimated hazard ratio and

associated 95% confidence intervals. Secondary analyses will be based on a Kaplan-Meier estimates on the all-randomized and per-protocol populations.

For the efficacy primary endpoint, the primary analysis will be based on a Kaplan-Meier analysis using the all-randomized analysis dataset. The 12-month event rates differences between the DOACs versus warfarin will be based on the difference in the Kaplan-Meier estimates and the one-sided upper 95% confidence interval will be based on a bootstrap estimator. Secondary analyses of the recurrent VTE endpoint will be based on the Cox regression model. Exploratory analyses will directly compare the two DOACs for both safety and efficacy endpoints. Additionally, an exploratory analysis will pool the two DOAC arms and compare those estimates with the warfarin treatment group.

For secondary endpoints, the estimated differences between the DOACs and warfarin treated groups will be based on linear models, logistic regression, and Cox models depending on the endpoint. Efforts will be made to limit the amount of missing data for the key endpoints and adherence measures. When necessary, multiple imputation will be used to account for missing covariate information.

References

Kaatz S, Ahmad D, Spyropoulos AC, Schulman S; Subcommittee on Control of Anticoagulation. Definition of clinically relevant non-major bleeding in studies of anticoagulants in atrial fibrillation and venous thromboembolic disease in non-surgical patients: communication from the SSC of the ISTH. J Thromb Haemost. 2015 Nov;13(11):2119-26.

Bates SM, Jaeschke R, Stevens SM et al. Diagnosis of DVT: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest 2012;141(2 Suppl):e351S-e418S.

Buckner TW, Key NS. Venous thrombosis in blacks. Circulation 2012;125(6):837-839.

Liao S, Woulfe T, Hyder S, Merriman E, Simpson D, Chunilal S. Incidence of venous thromboembolism in different ethnic groups: a regional direct comparison study. J Thromb Haemost 2014;12(2):214-219.

White RH, Keenan CR. Effects of race and ethnicity on the incidence of venous thromboembolism. Thromb Res 2009;123 Suppl 4:S11-S17.