Post Authorization Safety Study (PASS) Information

Acronym/Title	**PROTECT-AF**: A **P**ost-marketing **R**etrospective n**O**n-interventional study using na**T**ionwide registries and electronic medical records to investigate the real-life **E**ffectiveness and major bleeding **C**omplications of oral an**T**icoagulants in Norwegian non-valvular **A**trial **F**ibrillation patients.
Protocol version and date	v 1.0, 6 September 2018
IMPACT study number	19468
Study type / Study phase	Observational, Phase IV PASS Joint PASS: ☐ YES ☒ NO
EU PAS register number	Study not yet registered
Active substance	B01A F01 - Rivaroxaban ([redacted]) B01A F02 - Apixaban ([redacted]) B01A E07 - Dabigatran etexilate ([redacted]
Medicinal product	[redacted]
Product reference	EU/1/08/472/001-041
Procedure number	EMEA/H/C/00944
Comparator / Reference therapy	B01A A03 - Warfarin ([redacted]
Study Initiator and Funder	Bayer AG, 51368 Leverkusen
Research question and objectives	This population - based study will determine effectiveness and safety of oral anticoagulants in Norwegian non-valvular atrial fibrillation patients using the real-life data.
Country(-ies) of study	Norway
Author	[redacted]

	[redacted]

Marketing authorization holder

Marketing authorization holder(s)	[redacted]
MAH contact person	[redacted]

The study will be conducted in compliance with the protocol
and any applicable regulatory requirements.

Throughout this document, symbols indicating proprietary names (®, TM) may not be displayed. Hence, the appearance of product names without these symbols does not imply that these names are not protected.

1. Table of contents

2. List of abbreviations

AE	Adverse Event
ATC	Anatomical Therapeutic Chemical (Classification System)
AF	Atrial Fibrillation
CCI	Charlson Comorbidity Index
CI	Confidence Interval
DDD	Defined Daily Dose
DoD	Duration of Dispensation
DVT	Deep venous Thrombosis
eGFR	Estimated Glomerular Filtration Rate
EMA	European Medicine Agency
EMR	Electronic Medical Records
ENCePP	European Network of Centers in Pharmacoepidemiology and Pharmacovigilance
EU	European Union
FDA	Food and Drug Administration
FRG	The Norwegian Census Register
GPP	Good Publication Practice
GVP	Good Pharmacovigilance Practice
HR	Hazard Ratio
ICD	International Classification of Diseases
ICPC	International Classification in Primary Care
ICH	Intracranial haemorrage
IEC	Independent Ethics Committee
IRB	Institutional Review Board
INR	International Normalized Ratio
ISTH	Internation Society of Thrombosis and Hemostasis
MAH	Marketing Authorization Holder
N/A	Not Applicable
NCMP	Nordic Classification of Medical Procedures
NCSP	Nordic Classification of Surgical Procedures
NOAC	Non Vitamin K Antagonist Oral Anti-Coagulants
NOMESCO	The Nordic Medico-Statistical Committee
NorPD	Norwegian Prescription Database
NPR	Norwegian Patient Register
NVAF	Non-valvular atrial fibrillation
OAC	Oral Anti-Coagulants
OS	Observational Study
PAS	Post-Authorization Study
PASS	Post-Authorization Safety Study

PDC	Proportion Days Covered
PIN	Personal Identification Number
SAP	Statistical Analysis Plan
SPAF	Stroke prevention in atrial fibrillation
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
US	United States
VKA	Vitamin K antagonists
WHO DD	World Health Organization Drug Dictionary

3. Responsible parties

3.1 Study initiator and funder

Role: OS Conduct Responsible
Name: [redacted]
E-mail: [redacted]

Role: Qualified Person responsible for Pharmacovigilance (QPPV)
Name: [redacted]

Role: Regulatory Affairs
Name: [redacted]

Role: OS Safety Lead
Name: [redacted]

Role: OS Medical Expert
Name: [redacted]

Role: OS Statistician
Name: [redacted]

Role: OS Epidemiologist
Name: [redacted]

Role: OS Real Life Evidence Strategy and Outcomes Data Generation
Name: [redacted]

[redacted]

3.2 Collaborator(s)/External partner(s)/Committee(s)

[redacted]	Consultant Cardiologist	[redacted]	Principle Investigator
[redacted]	Senior Researcher	[redacted]	Project leader

4. Abstract

Acronym/Title	**PROTECT-AF**: A **P**ost-marketing **R**etrospective n**O**n-interventional study using na**T**ionwide registries and electronic medical records to investigate the real-life **E**ffectiveness and major bleeding **C**omplications of oral an**T**icoagulants in Norwegian non-valvular **A**trial **F**ibrillation patients.
Protocol version and date	v 0.1, 06 September 2018
IMPACT study number	19468
Study type / Study phase	Observational, Phase IV PASS
Author	[redacted]
Rationale and background	Non-vitamin K antagonist oral anticoagulants (NOACs) are approved for the prevention of stroke and systemic embolism in adult patients with non-valvular atrial fibrillation (NVAF) as they reduce the risk of these events. However, the anticoagulant effect of the drugs must be balanced against a risk of bleeding complications especially in patients with renal impairment and elderly who represent a large group of those affected by NVAF(1-3).
Research question and objectives	This population-based study will assess the effectiveness and safety of NOACs (rivaroxaban, dabigatran and apixaban compared to warfarin in OAC naïve NVAF patients in Norway. In addition, this study will serve as a pilot project to evaluate the ethical and operational feasibility as well as the usefulness of linking the EMR data to the Norwegian Health Registers. **Primary objective:** to assess the incidence rates of: a. ischemic stroke (effectiveness) b. intracranial hemorrhage (safety) in OAC naïve NVAF patients who are dispensed individual NOACs for the first time and compare with the corresponding

	rates in OAC naïve users of warfarin in a propensity score matched population.
	
	Secondary objectives
	• To assess the incidence rates of:
	  • overall stroke (ischemic, hemorrhagic, other unspecified) and systemic embolism outcomes;
	  • myocardial infarction;
	  • all-cause mortality;
	  • major bleeding as defined by ISTH and the Cunningham algorithm (in a sub-set of patients; cohort 2)
	in OAC naïve NVAF patients who are dispensed individual NOACs and compare with the corresponding rates in OAC naïve users of warfarin in the overall population and separately for the elderly and renally impaired patient using a propensity score matched population.
	• To describe and compare demographic and clinical characteristics (age, gender, CHA_2DS_2-VASc and HAS-BLED score, bleeding history, concomitant medications and co-morbidities) as well as drug utilization patterns of NVAF patients who initiated an OAC (NOACs and warfarin) in the study period.
	
	Several exploratory objectives aim to assess:
	  • differences in characteristics of NVAF patients receiving OACs or not
	  • impact of adherence to NOACs on rates of ischemic stroke and bleeding
	  • patient characteristics of NVAF patients receiving standard vs. reduced dose of NOACs and corresponding rates of ischemic stroke and bleeding in these sub-groups
	  • validate data on indications and outcomes derived from the national registers using Electronic Medical Records

	(EMR) available for a regional sub-set of patients (cohort 2).
Study design	This study will be a population-based retrospective cohort study on de-identified individual-level data extracted from national registers and EMRs in Norway. It will be based on two partly overlapping cohorts of patients, one utilizing only nation-wide health registers (cohort 1) and one cohort based on EMR data from selected hospitals in Norway combined with the registry data (cohort 2).
Population	The study population will comprise all adult OAC naïve NVAF patients in Norway who filled a prescription for an OAC (rivaroxaban, apixaban, dabigatran and warfarin) in the study period, defined as from 1 January 2014 to 30 June 2018 (or later depending on availability of data) and followed until outcome of interest, the end of the study period or death. This initiation of an OAC is the index event and requires that there is no previous prescription dispensed for an OAC in the preceding 365 days (counted from the date of the first dispensation backward to the end of OAC supply).
Variables	Patient demographics (age, gender, migration status, place of birth and date of death), clinical characteristics on outcomes, co-morbidities (ICD10 codes and procedure codes with date for admission and discharge) and prescription drugs used both OAC and relevant co-medication (ATC codes with date of dispensation, pack size, strength, number of units and so forth) will be retrieved from the nationwide registries. Risk scores (HAS-BLED and CHA2DS2-VASc) will be calculated. More detailed demographic and clinical data such height, weight, smoking and alcohol use, laboratory values (INR, hemoglobin, s-creatinine) and dosing of OAC will be retrieved from the regional EMR data.
Data sources	De-identified individual level data from three nation-wide registers will form the basis for the analyses, the Norwegian Census Register (FRG), the Norwegian Prescription Database (NorPD) and the Norwegian Patient Register (NPR).

	Additionally for cohort 2, in the South-East region of Norway, an automated EMR extraction tool will be deployed in order to retrieve additional clinical data that is not present in the nationwide registers.
Study size	The study population will comprise all adult OAC naïve NVAF patients in Norway who filled a prescription for an OAC (rivaroxaban, apixaban, dabigatran, warfarin) in the study period, defined as from 1 January 2014 to 30 June 2018 (or later depending on availability of data). Based on the information regarding number of patients receiving listed drugs in the country, the sample size for the study will be approximately 70,000 patients.
Data analysis	To compare warfarin with apixaban, dabigatran and rivaroxaban the incidence rates and corresponding 95% confidence intervals for ischemic stroke and intracranial hemorrhage will be calculated in a propensity score matched population. Additionally the incidence rates of: • overall stroke and systemic embolism outcomes; • myocardial infarction; • all-cause mortality; • major bleeding will be calculated.
Milestones	Protocol submission to IEC: Q3 2018 IEC approval: Q4 2018 Data inspectorate and registry holder approvals: Q1 2019 Analytical dataset available Q2: 2019 Preliminary results available on primary objective: Q2 2019 Final and complete report: Q1 2020

5. Amendments

None

6. Milestones

Table 1: Milestones

Milestone	Planned date
Protocol submission to PRC	5 April 2018
Ethical Committee approval	30 October 2018*
Registry holder approval	15 January 2019
Start of data collection	15 January 2019
End of data collection	15 March 2019
Analytical dataset available	15 April 2019
Preliminary data on primary objective	15 June 2019
Data Analysis Completed	15 October 2019
Registration in the EU PAS register	
Final report of study results	15 January 2020

*assuming one round of comments from Ethical committee

7. Rationale and background

NOACs are approved for prevention of stroke and systemic embolism in adult patients with non-valvular atrial fibrillation (stroke prevention in atrial fibrillation [SPAF]) as they reduce the risk of clot formation. However, the anticoagulant effect of the drugs must be balanced against a risk of bleeding complications especially in patients with renal impairment and elderly who represent a large group of those affected by Atrial Fibrillation (AF)(1-3).

In Norway, all Non vitamin K antagonist Oral Anti-Coagulants (NOACs) received full reimbursement for SPAF relatively soon after launch. There is no difference in reimbursement conditions between the different NOACs and they have been widely used in several indications including SPAF.

A recently published study based on the Norwegian population registry data indicated different incidence rates of bleeding for the different NOACs compared to warfarin.(4) The same study identified a significant proportion of patients receiving the reduced dose of NOACs, the

appropriateness however of the reduced dosing could not be assessed due to the lack of clinical information. Another clear limitation of this study is the lack of effectiveness data and information on adherence.

Norway has a relatively homogenous and stable population of just over 5 million and a publicly funded healthcare system. Health-related data is systematically captured in electronic format in several registers, including National Health Registers that contain data on hospitalizations, prescriptions as well as causes of death of all inhabitants. Hospital and primary care Electronic Medical Records (EMRs) information are also available. Inhabitants of Norway have a unique personal identification number (PIN) which is used to link information from the various health registers.

Extraction of large amounts of data from EMRs is often a limiting factor. An automated data extraction tool available in Scandinavia (Customised eXtraction Program [CXP], IQVIA) was shown to efficiently and reliably extract EMR data from hospitals and primary care clinics across the country in full adherence to the data protection rules(5). This unique methodology automatically extracts structured and unstructured EMR data for large numbers of patients and assigns identifiers that make cross-linking with registry data possible on an individual patient level, while ensuring patient anonymity. The methodology of the CXP tool was developed in Sweden and it has been applied in primary care in Norway. However it has not yet been applied with Norwegian hospital EMRs.

8. Research questions and objectives

This population-based study offers a unique opportunity to assess both the effectiveness and safety of NOACs (rivaroxaban, dabigatran, apixaban) compared to warfarin in OAC naïve NVAF patients in Norway. In addition, this study will serve as a pilot project to evaluate the operational feasibility as well as the usefulness of linking the EMR data to the Norwegian Health Registers.

8.1 Primary objective:

- To assess the incidence rates of:
    a. ischemic stroke (effectiveness)
    b. intracranial hemorrhage (safety)

in OAC naïve NVAF patients who are dispensed individual NOACs for the first time and compare with the corresponding rates in OAC naïve users of warfarin in a propensity score matched population.

8.2 Secondary objectives

- To assess the incidence rates of:

- overall stroke (ischemic, hemorrhagic, other unspecified) and systemic embolism outcomes;
- myocardial infarction;
- all-cause mortality;
- major bleeding as defined by ISTH and the Cunningham algorithm (in a sub-set of patients; cohort 2)

in OAC naïve NVAF patients who are dispensed individual NOACs for the first time and compare with the corresponding rates in OAC naïve users of warfarin in the overall and in a propensity score matched population.

- To describe and compare demographic and clinical characteristics (age, gender, CHA_2DS_2-VASc and HAS-BLED score, bleeding history, concomitant medications and co-morbidities) as well as drug utilization patterns of NVAF patients who initiated an OAC (NOACs and warfarin) in the study period.

8.3 Exploratory objectives (naïve OAC users)

- To compare demographic and clinical characteristics (age, gender, CHA_2DS_2-VASc and HAS-BLED score, bleeding history, concomitant medications and co-morbidities) and the incidence rates for ischemic stroke and intracranial hemorrhage in AF population that is not treated with OACs with AF population being treated with OACs.
- To investigate sub-groups of interest (patients > 75 years, renal impairment) in stratified analyses.
- To describe and compare adherence levels to NOACs among NVAF patients and assess whether different levels of adherence is associated with changes in the incidence rate of ischemic stroke and bleeding outcomes.
- To describe patient characteristics of NVAF patients receiving standard vs. reduced dose of NOACs and to estimate corresponding rates of ischemic stroke and bleeding
- To validate the diagnosis of NVAF, bleeding and stroke endpoints derived from the NPR data using EMR data.

9. Research methods

9.1 Study design

This study will be a population-based retrospective cohort study on de-identified individual-level data extracted from registers and electronic medical records (EMRs) in Norway. It will be based on two partly overlapping cohorts of patients, one utilizing only nation-wide health

registers (called cohort 1) and one cohort based on EMR data from selected hospitals in Norway combined with registry data (cohort 2).

Warfarin has been the traditional treatment in Norway for many years, while dabigatran was marketed for NVAF patients from 1 July 2008, rivaroxaban from 1 February 2012 and apixaban from 15 February 2013. The three NOACs were granted Norwegian general reimbursement from 1 January 2013 (dabigatran and rivaroxaban) and 15 July 2013 (apixaban). The study period (i.e. from 1 January 2014) was defined to incorporate all use of the NOACs once they all had been under general reimbursement for at least 6 months.

Table 2: OAC marketing authorization dates and reimbursement dates for NVAF indication[1]

ATC code	Drug	Marketed	NVAF Reimbursement
B01AA03	warfarin	4 October 1962	Early
B01AF02	apixaban	15 February 2013	15 July 2013
B01AF01	rivaroxaban	1 February 2012	1 January 2013
B01AE07	dabigatran	1 July 2008	1 January 2013

9.2 Setting

The study will be undertaken in Norway and will comprise all adult patients with NVAF who initiated treatment with either a NOAC or warfarin during the study period for the first time (OAC naïve first-time users). Study endpoints, discontinuation, switch to another OAC, death or end of study period will serve as censoring criteria.

The NVAF study population is defined in accordance with the 2018 European Heart Rhythm Association practical guide on the use of non-vitamin K antagonist oral anticoagulants in patients with atrial fibrillation, which define NOAC eligibility in NVAF patients as atrial fibrillation in the absence of rheumatic mitral stenosis or a mechanical heart valve.

Cohort 1 (registry study): Registries in Norway are nation-wide and provision of the information is mandatory, which eliminates the risk of both selection and re-call bias. The large and detailed dataset also makes it possible to adjust for other risk factors on which information is available.

In cohort 2, patients with NVAF diagnosis will be identified through extraction of patient-level data from EMRs from a number of hospitals in Norway, in order to describe these patients more closely regarding their clinical characteristics that are not available in nation-wide registers (e.g., in-patient treatments, anthropometric data and laboratory test results). This part of the

[1] Source: http/www.legemiddelverket.no

study is a pilot that will test the operational feasibility of accessing and linking the regional EMR data to the Norwegian Health Registers. Should the access to the regional EMR be granted, patients in cohort 2 will be identified and their EMRs extracted using the Pygargus Customized Extraction Platform (CXP). The extraction program has been used in a number of research projects since 2005 in the Swedish and Norwegian health care systems (6, 7). De-identified individual level data on age, gender, prescriptions, diagnoses, measures, laboratory tests, health care professional (HCP) visits, referrals, etc. can be collected from the EMRs into an encrypted database.

A key-code linked to each individual patient ID number will be extracted to enable linkage to national healthcare registers. This data will then be linked with the national registers data used in cohort 1. This key-code will be stored separate from the database used for analysis. Individual patient data from EMRs will then be linked to the described national registers (Norwegian Patient Registry, Norwegian Prescription Database and the Cause of Death Register). Extraction of patient level data will take place once and cover the period from 2009 to 2017 (or later depending on when EMR extraction can take place).

Datasets for cohort 1 and 2 will be de-identified before they are accessible for analysis. Neither Bayer nor other research partners will have access to patient-identifiable data. This is a legal prerequisite and is strictly enforced by the Registry Holders in Norway and is thoroughly reviewed during the application for an ethics approval. Furthermore, it is planned to store and analyze the datasets within the secure environment of "Services for sensitive data (TSD)" at the University of Oslo. TSD offers a secure environment for the collection, storage and analysis of sensitive research data.

9.2.1 Study time frame

For the main analysis, the study population will comprise all adult OAC naïve (no OAC prescription in the previous 365 days, counted from the date of the first dispensation backward to the end of OAC supply) NVAF patients in Norway dispensed a prescription for an OAC (NOAC or warfarin) for the first time in the study period (cohort 1a), defined as from 1 January 2014 to 30 June 2018 (or later depending on availability of data) and followed until endpoint of interest, the end of the study period or death. Patients will be censored at a switch or a discontinuation.

For exploratory analysis, a second study population (cohort 1b) will comprise all adult NVAF patients in Norway that did not receive an OAC (NOAC or warfarin) in the study period, defined as above.

The look-back period will be at least 5 years from the possible earliest index date of NOACs (1 January 2014) going back to 1 January 2009 (available data cut). The minimum follow-up period will be 1 month and the estimated mean follow-up time will be 18 months.

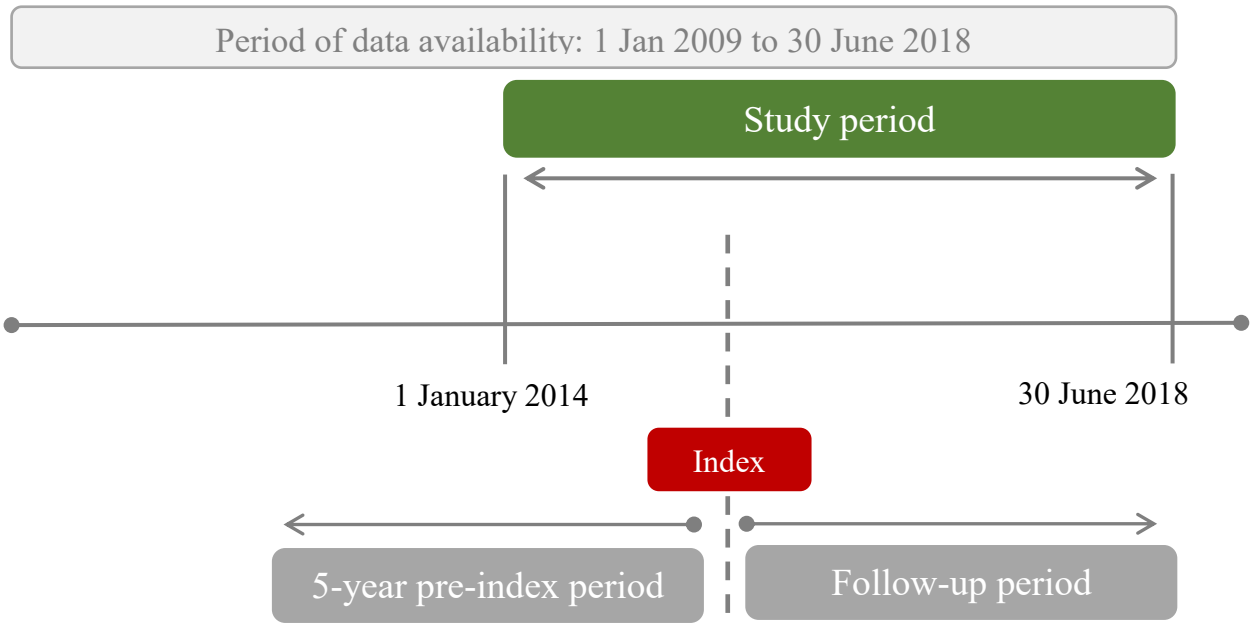

Figure 1: Study design

9.2.2 Selection criteria

Depending on the data source, the following 2 cohorts will be recruited:

Inclusion criteria
Registry cohort (Cohort 1)

Cohort 1a:

- Age ≥18 at the date of first OAC dispensation (index date)

- Diagnosed with atrial fibrillation or flutter in the period 5-year pre-index period. Defined as ICD10 codes; I480 (paroxysmal AF), I481 (persisting AF), I482 (chronic AF), I483 (typical AF), I484 (atypical AF) or I489 (unspecified atrial fibrillation or flutter) as given by either NPR **or** as a OAC (warfarin, dabigatran, rivaroxaban or apixaban) dispensed on the reimbursement code for AF in NorPD (ICD10 I48 or ICPC-2 K78)

- Dispensation of a first-time prescription for an OAC during the study period 1 January 2014 to 31 May 2018. This initiation of an OAC is the index event and requires that there is no previous prescription for any OAC in the preceding 365 days (counted from the date of the first dispensation backward to the end of OAC supply).

Cohort 1b:

- Age ≥18 at the beginning of the study period (1 January 2014).

- Diagnosed with atrial fibrillation or flutter prior or during the study period. Defined as ICD10 codes; I480 (paroxysmal AF), I481 (persisting AF), I482 (chronic AF), I483 (typical AF), I484 (atypical AF) or I489 (unspecified atrial fibrillation or flutter) as given by NPR**.**

- No dispensation of a prescription for an OAC (warfarin, dabigatran, rivaroxaban or apixaban) during the study period 1 January 2014 to 30 June 2018. The index event will be the beginning of the study period (1 January 2014) and requires that there is no previous prescription for an OAC in the preceding 365 days.

EMR cohort (Cohort 2):

- Age ≥18 at the date of first OAC dispensation (index date).

- Diagnosed with atrial fibrillation or flutter in the 5-year pre-index period defined as one of the following ICD10 codes in the EMRs obtained from the hospitals in the South-

Eastern Health Region; I480 (paroxysmal AF), I481 (persisting AF), I482 (chronic AF), I483 (typical AF), I484 (atypical AF) or I489 (unspecified atrial fibrillation or flutter).

- Dispensation of a first-time prescription for an OAC (warfarin, dabigatran, rivaroxaban or apixaban) during the study period 1 January 2014 to 31 May 2018. This initiation of an OAC is the index event and requires that there is no previous prescription of any OAC in the preceding 365 days (counted from the date of the first dispensation backward to the end of OAC supply).

Exclusion criteria (for both cohorts)

- Valvular atrial fibrillation defined as presence of ICD10 codes in NPR; I05.2 mitral stenosis with insufficiency, I05.8 other mitral valve diseases (mitral (valve) failure), I342 (non-rheumatic mitral valve stenosis), Z952 (presence of prosthetic heart valve) and procedure codes FKD (prosthetic replacement of mitral valve), FKA (Repair of mitral valve for stenosis), FKB (Annuloplasty of mitral valve for insufficiency), FKC (Repair of mitral valve for insufficiency), FKW (Other operations on mitral valve) and FMD (replacement of aortic valve).
- Deep Venous Thrombosis (DVT) as defined by ICD10 codes I80 (Phlebitis and thrombophlebitis), Pulmonary Embolism (PE) as defined by ICD10 code I26 or I82 (other venous embolism and thrombosis), Other venous embolism and thrombosis within last 6 months before index date.
- Knee and/or hip replacement surgery (NCSP procedure codes; NGB, NGC, NFB or NFC) 5 weeks before index date.

9.2.3 Study population

Cohort 1 sampling:

The NorPD is a pseudonymized register which means that it is not possible to use dispensation of OAC as the starting point and as a source population. Being a psedonymized register NorPD data needs to be the last register in the data collection that is linked to the other data sources. To circumvent this we will start with the Norwegian Census Register that will identify all Norwegian residents' of the age of 18 years and above during the study period. This dataset is sent to the NPR (for all patients with NVAF and without NVAF plus the relevant co-morbidities). Lastly, NorPD will identify all subjects with at least one OAC dispensation during the study period and link the registers (cohort 1a). A de-identified dataset with all NVAF patients having at least one OAC dispensed in the study period will be delivered to the research group (cohort 1a). Similarly, NorPD will identify all subjects that did not receive an OAC prescription during the study period and link the registers (cohort 1b). A de-identified dataset with all NVAF patients not having received an OAC prescription in the study period will be delivered to the research group (cohort 1b) for exploratory analysis.

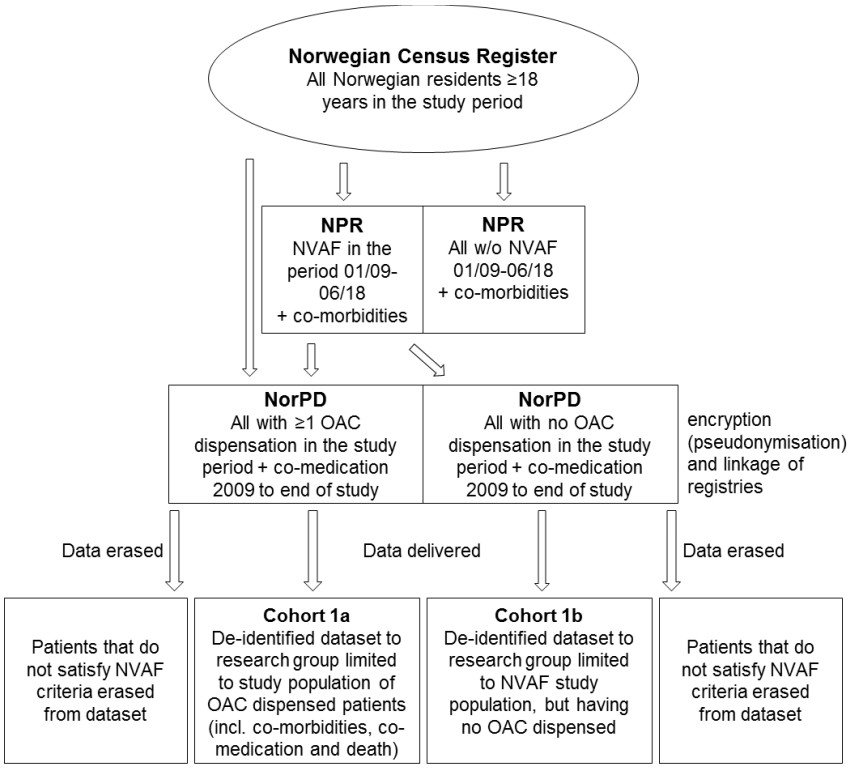

Figure 2: Cohort 1 sampling procedure flow-chart

Cohort 2 sampling:

From the EMR extraction process all patients diagnosed with NVAF at selected hospitals in Norway between 2009 to the end of study will be selected. This cohort of patients will be linked with NPR and lastly NorPD will be added.

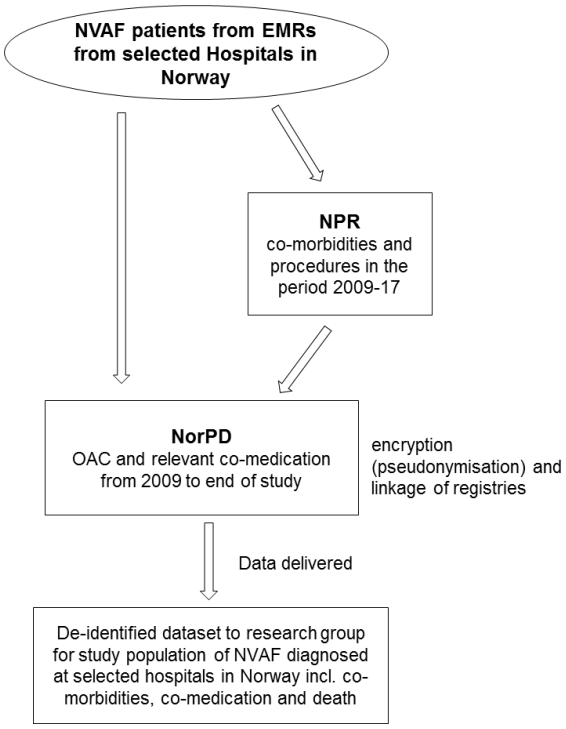

Figure 3: Cohort 2 sampling procedure flow-chart

9.2.4 Representativeness

The study captures the whole Norwegian population, ensuring representativeness for the country. Moreover, the Prescribed Drug Register is-almost complete. It keeps track of every single pill that is dispensed, in any pharmacy throughout the country, and this information is linked to the individual who received it. The quality of the Norwegian health registers is well known and they provide a unique source of information for post-marketing studies of drug use and outcomes.

9.3 Variables

9.3.1 Baseline characteristics

Table 4: Demography and clinical characteristics extracted for cohort 1 and 2

Variable	**data source**	**Definition**	**Relevance**
Age	FRG	Age in year	Co-variate and included in risk scores
Gender	FRG	Male or female	Co-variate and included in risk scores
Origin	FRG	Place of birth categorized as Norway, the Nordic countries, other Europe, rest of the world	Co-variate
Immigration date	FRG	Date of immigration out of Norway	Censoring
Emigration date	FRG	Date of emigration out of Norway	
Date of death	FRG	Date of death	Censoring and outcome variable
Alcoholism	NPR	ICD10 codes: E244, G312, G621, G721, I426, K860, O354, Z714, Z721, E52, K70, T51	Co-variate
Chronic kidney disease (CKD)	NPR	ICD10 code: N18 as an option 1 to define CKD Expanded list of codes to define CKD will be considered as an option 2, This list is being used for another ongoing study and is based on the approaches by Fleet et al (8)., Nielsen et al (9). and Weir et al (10). that in addition to identifying patients from the option 1 will allow an inclusion of patients with cystic kidney disease, unspecified kidney failure, chronic or unspecified nephritic syndrome, nephrotic syndrome, recurrent and	Co-variate

Variable	data source	Definition	Relevance
		persistent hematuria, nephropathy (diabetic, hypertensive, hereditary), chronic tubulo-interstitial nephritis, diabetes mellitus with kidney complications as having renal dysfunction (see Annex 3). The option 2 will be further clarified.	
End stage renal disease	NPR	ICD10 code: N185 or procedure codes KAGD40 (hemodialysis chronic) or KAS (renal transplant)	Co-variate Hemodialysis included in mHASBLED score
Congestive heart failure	NorPD	NorPD reimbursement codes for heart failure: I50 (ICD10) or K77 (ICPC) in ATC code group C	Co-variate. Included in CHA2DS2VASc score
Dementia	NPR	ICD10 codes F00-F04, G30	Co-variate
Diabetes	NorPD	ATC code A10A or A10B	Co-variate
Ischemic stroke	NPR	ICD10 codes: I63	Co-variate
Hemorraghic stroke	NPR	ICD10: I60-62	Co-variate Included in mHAS-BLED score
TIA	NPR	ICD10 code: G45	Co-variate Included in CHA2DS2VASc score
Stroke&tia	NPR	ICD10: I60-I64	Co-variate Included in CHA2DS2VASc score
Thromoembolism	NPR	ICD10: I74	Co-variate

Variable	data source	Definition	Relevance
			Included in CHA2DS2VASc score
Hypertension	NorPD	Reimbursement codes in NorPD in ATC code group C: E10-E13, E15 (ICD10) or K86, K87 (ICPC)	Co-variate Included in mHASBLED and CHA2DS2VASc risk score
Peripheral artery disease	NPR	ICD10: I739	Co-variate Included in CHA2DS2VASc score
Ischemic heart disease	NPR	ICD10: I20-I25	Co-variate
Myocardial infarction	NPR	ICD10: I21-I22	Co-variate
Angina pectoris	NPR	ICD10: I20	Co-variate
Chronic obstructive pulmonary disease	NorPD	Reimbursement codes: J44 (ICD10) or R95 (ICPC) in ATC group R03	Co-variate
Cancer	NPR	ICD10 codes: C00 to C97	Co-variate Included in Charlson Co-morbidity score
Leukemia	NPR	ICD10 codes: C90-C96	Co-variate Included in Charlson Co-morbidity score
Lymphoma	NPR	ICD10 codes: C80 to C88	Co-variate Included in Charlson Co-morbidity score
AIDS	NPR	ICD10 codes B20 to B24	Co-variate

Variable	data source	Definition	Relevance
			Included in Charlson Co-morbidity score
Thrombocytopenia	NPR	ICD10 codes: D692, D693, D694, D695, D696, D698, D699	Co-variate
Anemia	NPR	ICD10 codes: D65-D67, D680, D681, D682, D683, D684, D685, D686, D688, D689, D50-D53, D55-D59, D60-D64	Co-variate
Hypo/hyperthyriodism	NPR	ICD10 code: E03 / E05	Co-variate
Systemic connective tissue disease	NPR	ICD10 codes: M30 to M36	Co-variate Included in Charlson Co-morbidity score
Peptic ulcer and bleeding	NPR	ICD10 codes: Ulcer K25 to K29. Gastrointestinal bleeding K920/K921/K922	Co-variate Included in Charlson Co-morbidity score
Chronic liver disease and impairment	NPR	ICD10 codes: K70-K77	Co-variate Included in mHASBLED score
Renal impairment	NPR	ICD10 codes: D63.8, E10.2, E11.2, E13.2, I12.0, I13.1, I13.2, N02.x, N03.x, N04.x, N05.x, N07.x, N08.x, N11.x, N14.x, N18.x, N19, Q61.x	Co-variate Used in stratified analyses
Surgery performed in the study period	NPR	All NCSP codes including codes for active surgery	Co-variate. Adjust or exclude for in bleeding outcomes
Imidazol and triazol derivates	NorPD	ATC codes J02AB and J02AC	Co-variate. Interaction

Variable	**data source**	**Definition**	**Relevance**
Low-dose aspirin	NorPD	ATC code: B01A C06	Co-variate Included in mHASBLED score
Immunosuppressive treatments	NorPD	ATC code: L04A	Co-variate
Anti-platelet treatment (non aspirin)	NorPD	ATC code: B01A	Co-variate Included in mHASBLED score
Non-steroidal anti-inflammatory drugs (NSAID)	NorPD	ATC code: M01A	Co-variate. Included in mHASBLED score
Acid secretory drugs	NorPD	ATC code: A02B	Co-variate
Heparin	NorPD	ATC code: B01AB	Co-variate
Anti-arrytmic drugs class iii	NorPD	ATC code C01BD	Co-variate. Interaction
Rifampicin	NorPD	ATC code J04AB	Co-variate. Interaction
Anti-hypertensives	NorPD	ATC code: C02	Co-variate
Diuretics	NorPD	ATC code: C03	Co-variate
Beta-blockers	NorPD	ATC-code: C07	Co-variate
Calcium antagonists	NorPD	ATC-code: C08	Co-variate
Renin-angiotensin system drugs	NorPD	ATC-code: C09	Co-variate

Variable	data source	Definition	Relevance
Lipid-modifying drugs	NorPD	ATC-code: C10	Co-variate
Reduced or standard dose NOAC at index	NorPC	Indicator variable whether on reduced (15 mg riva, 2,5 mg api, 110 mg dabi) or standard (20 mg riva, 5 mg api, 150mg dabi) at index dispensation	Co-variate
Prior OAC use	NorPD	Previous OAC use. OAC availability (time of exhaustion of last dispensation) in 365 days prior to index dispensation. From the time of exhaustion of last dispensation and initiation of OAC in study period (> 12 months pre-index)	Co-variate
Previous bleeding hospitalization	NPR	ICD codes listed in section 9.3.4.2	Co-variate Included in mHASBLED
Modified HAS-BLED score	NPR and NorPD	− Hypertension defined by dispensation of at least one anti-hypertensive drug (see definition for hypertension above) last year prior to index date. 1 point − Renal impairment as defined by either hemodialysis (procedure code KAGD40) or kidney transplantation (procedure code KAS) or CKD stage 3 and 4 (ICD10 code N183, N184, N185). Expanded list of codes to define CKD will be considered as an option 2, 1 point − Liver impairment (ICD10 K70-K77). 1 point − Stroke during the last year prior to index visit (see definition). 1 point − Prior major bleeding (critical organ) last year before index date (see definition in 9.3.4.2) 1 point	Bleeding risk score used as co-variate and effect modifier

Variable	data source	Definition	Relevance
		– Age ≥ 65 years. 1 point – Therapy with either Nonsteroidal Anti-inflammatory Drugs (NSAID M01A) or anti-platelets (B01AC) in the previous year before index date. 1 point – Alcoholism (as defined by ICD10 codes above). 1 point Total HAS-BLED score from 0 to 8	
CHA_2DS_2VASc	NPR and NorPD	- Chronic Heart Failure by NorPD reimbursement codes for heart failure: I50 (ICD10) or K77 (ICPC). 1 point - Hypertension defined by dispensation of at least one anti-hypertensive drug last year prior to index date (see definition above). 1 point - Diabetes defined by dispensation of at least one anti-diabetes drug last year prior to index date (see definition above). 1 point - Age ≥75 years (2 points) or age 64-74 years (1 point) - Previous stroke, TIA or thromboembolism by ICD10 codes (see definition above). 2 points - Peripheral artery disease (ICD10; I739). 1 point - Gender (male=1 point, female=0 point) Total score 0 to 9	Stroke risk score used as co-variate and effect modifier
Charlson co-morbidity index	NPR and NorPD	- Age: ≤41 years (0 point), 41-50 years (1 point), 51-60 years (2 points), 61-70 (3 points), ≥71 years (4 points) - Myocardial infarction. 1 point - Congestive Heart Failure. 1 point - Peripheral Vascular Disease. 1 point - TIA or stroke. 1 points - Dementia. 1 point	Co-morbidity risk score and co-variate

Variable	data source	Definition	Relevance
		- COPD. 1 point - Connective Tissue Disease. 1 point - Peptic ulcer and bleeding 1 point - Diabetes. 1 point - Moderate to severe Chronic Kidney Disease (combined CKD and ESRD). 2 points - Leukemia. 2 points - Lymphoma. 2 points - Solid tumour. 2 points - Liver disease. 3 points - Solid tumour with metastasis. 6 points - AIDS. 6 points Total score from 0 to 37	

Table 5: Additional clinical characteristics to be retrieved for the EMR cohort (cohort 2)

Variable	**Data source**	**Operation definition**	**Relevance**
Height	EMR	Height in centimeters from time point closest to the index date	Co-variate
Weight	EMR	Weight in kilogram from time point closest to the index date	Co-variate
Smoking habits	EMR	Smoking habits retrieved as free text search. Converted to pack-years	Co-variate
Alcohol abuse	EMR	Alcohol abuse search as free text	Co-variate
International normalised ratio	EMR	Retrieve all INR measurements from index to the end of study. Calculated INR mean during the study period	Co-variate. Imputation
Blood haemoglobin	EMR	Retrieve all Hb measurements from index to the end of study. Calculated Hb mean during the study period. Identify episodes of more than 20 g/L fall in hemoglobin	Co-variate. Bleeding endpoint
S-Creatinine	EMR	Retrieve all s-creatinine measurement from index to the end of study. Calculate eGFR (ml/min) and categorise in CKD stages (1-5)	Co-variate
OAC dosing	EMR	Number of tablets used per days retrieved from dosing instructions. Especially important for warfarin dosing	Co-variate. Imputation and modelling of warfarin exposure in the registry cohort
Blood transfusions	EMR	Number of blood transfusion. More or equal to two transfusion units will be used in the ISTH bleeding definition	For ISTH bleeding outcome

Variable	Data source	Operation definition	Relevance
Fatal bleeding	EMR	Patients that die from bleeding while at hospital	For ISTH bleeding outcome

9.3.2 OAC exposure

New OAC exposure is defined as a single or series of dispensation of an OAC starting in the period from 1 January 2014 to end of study period. The duration of the observation period together with the pack size of each prescription will be used to calculate the length of a patient's OAC exposure. To ensure a minimum of one month follow-up data for all subjects on OAC treatment, inclusion (cut-off index OAC date) will be stopped one month before the end of available follow-up data.

For each dispensation, the OAC days of supply will be computed using information on date of dispensation, the number of packages and the pack-size dispensed. Since NOACs are prescribed in a fixed dose, the number of days of supply strictly corresponds to amount dispensed accounting for a recommended frequency of daily dosing. The NorPD contains information on tablet strength, pack-size and number of packages dispensed, and we assumed, according to the labelling, twice daily dosing for apixaban and dabigatran and once daily dosing for rivaroxaban, e.g. a patient supplied one package of a 100 tablet package of rivaroxaban will have a supply lasting for 100 days, whereas a 100 tablet package of apixaban will provide a supply lasting 50 days.

Computing the warfarin supply is not straightforward as we lack information on both dosing instructions and International Normalised Reference values (INR) in the registries.

Two different methods will be applied to compute the warfarin exposure;

1. EMR data (from cohort 2) holds information on both INR and dosing. Median mg/day of warfarin will be computed in relation to age (in 5-year categories) and gender (females and males). The median mg/day will be used as a basis for computing the warfarin supply in cohort 1.
2. Registry data (from cohort 1) using the length of each time interval in days between dispensation and corresponding dispensed warfarin amount for the time interval will be used to compute the median daily warfarin use for each patient. The calculated median mg/day for all patients using warfarin in the study period, in relation to age group and gender as with method 1, will subsequently be used in the computation of warfarin supply for each dispensation.

Similarly, the same methods will be used to set the end of OAC supply date during the pre-index period to be able to determine if a patient was OAC naïve or not (≥ 365 days without OAC supply prior to index date).

The duration of dispensation (DoD) will be set as the date of a dispensed prescription plus the estimated duration of that prescription (duration of OAC supply estimated from the dispensed pack-size and strength).

The following rules applies:

1. If the next dispensed prescription was for the same OAC within 30 days after end of DoD there was a continuous treatment
2. If a new dispensed prescription for another OAC occurred within DoD or within 30 days after end of DoD there was a switch in OACs (switching date = date of dispensation)
3. If the next dispensed prescription was for the same OAC more than 30 days after end DoD there was a discontinuation (discontinuation date = date of DoD + 30 days)

A 30-day gap period is included to take into account incomplete adherence to treatment. The figure underneath displays the three different scenarios. In sensitivity analyses the gap period will be varied from 30 days to 0 days, 15 days and 60 days, respectively.

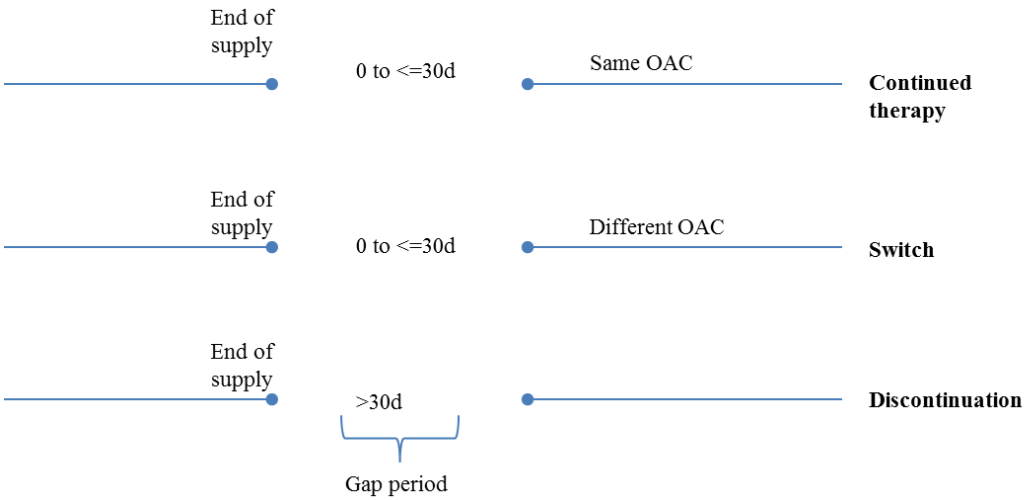

Figure 4: Definition of continuation, switch and discontinuation

Patients to be included in the effectiveness and safety analysis should have a supply of OAC covering the event, plus a 30 days grace period for those patients discontinuing their therapy. Any event happening after the date of discontinuation plus 30 days or after a switch in OAC therapy will not be included.

9.3.3 Estimation of OAC adherence

Proportion of days covered (PDC) will be estimated for each individual from the index medication date and onwards. The PDC is defined as the number of days that the index medication was on hand during the time period divided by the total number of days in the period. For instance if the length of the time period is 100 days and the NOAC supply covers 50 days, the adherence is estimated to 50%.

Likewise, the age- and sex-specific median doses of warfarin (using the two methods described in 9.3.2) will be used to calculate to warfarin supply.

If the OAC supply last longer than the study period the adherence is set to 100%. The proportion of days covered (PDC) will take into account the patient's entire follow-up period, defining end of study period as death, end of study, migration, discontinuation or switch of OAC therapy. In addition to the entire period, adherence will also be estimated at fixed time intervals after OAC index date, i.e. 3 months (90 days), 6 months (180 days), and 9 months (270 days), 1 year (365 days), 2 years (730 days) and 3 years (1095 days).

Consecutive dispensations of the same OAC will be combined and corrected for overlap. For treatment gaps equal or less than 30 days of the previous OAC dispensation it is assumed the patient is on continuous therapy. The length of the gap period will be varied in sensitivity analyses down to 0 and 15 days and up to 60 days.

Some patients switch to another NOAC during follow-up and the day of switch will be defined as the end of the time period. Separately, the PDC for use of any OAC during the study period will be estimated, which accounts for therapy switching during follow-up.

PDC will be estimated as the mean PDC for each OAC as well as grouped into four categories of high adherence (≥ 0.80), moderate adherence (0.50-0.79), low adherence (0.20-0.49) and poor adherence (<0.20). Adherence will also be reported as dichotomized using the conventional PDC of <0.80 or ≥ 0.80 as cutoff. Percentage point differences between OAC user groups will be presented. Analyses of OAC adherence will account for differences in follow-up time and baseline characteristics, such as standard and reduced dosing level.

9.3.4 Outcome measures

9.3.4.1 Outcome definitions effectiveness

Cohort 1 and 2 (same for both cohorts):

The main efficacy endpoint is ischemic stroke defined by the following ICD10 codes:

I63x:	Cerebral infarction

Overall stroke is defined by the following ICD10 codes:

I60x:	Subarachnoid haemorrhage
I61x:	Intracerebral haemorrhage
I62x:	Other nontraumatic intracranial haemorrhage
I63x:	Cerebral infarction
I64x:	Stroke, not specified as haemorrhage or infarction

Events of systemic embolism will be analysed separately and in combination with ischemic and overall stroke. Systemic embolism is defined by the following ICD10 codes:

I74x: Arterial embolism and thrombosis

All-cause mortality will be analysed separately and in combination with ischemic and overall stroke plus systemic embolism.

Events of myocardial infarction will be analysed separately and is defined by the following ICD10 codes:

I21x: Acute myocardial infarction
I22x: Subsequent myocardial infarction

9.3.4.2 Safety outcome definitions

Cohort 1:

NPR lacks information on laboratory measurements. Hence, we cannot use the ISTH (International Society of Thrombosis and Hemostasis) definition of a major bleeding episode which include a fall in hemoglobin level of 20 g/L (1.24 mmol/L) or more

The primary bleeding endpoint of intracranial haemorrage (ICH) is defined by the following ICD10 codes:

I60x Subarachnoid haemorrhage
I61x Intracerebral haemorrhage
I62x Other non-traumatic intracranial haemorrhage

Cohort 2:

The primary bleeding endpoint of intracranial haemorrage (ICH) is defined by the following ICD10 codes:

I60x Subarachnoid haemorrhage
I61x Intracerebral haemorrhage
I62x Other non-traumatic intracranial haemorrhage

Secondary safety outcome measures in cohort 2 are major bleeds defined by the Cunningham algorithm (11) and the ISTH definition (12) for major bleeds.

The Cunningham algorithm (11) defines a hospitalization with a primary ICD10 code for major bleeding except for ICH where all hospitalizations is counted (> 8 hrs in-patient)(see below).

The Cunningham algorithm is originally based on the ICD-9-CM system. All codes in the ICD-9-CM is not easily converted to ICD9 and later the ICD9 system was replaced by ICD10. Thus, the ICD10 based Cunningham definition will be built on the following outcomes:

Gastrointestinal:

K920	Hematemesis
K921	Melena
K922	Gastrointestinal haemorrhage, unspecified
K226	Mallory-Weiss syndrome
K250/K252/K254/K256	Gastric ulcer with bleeding and/or perforation
K260/K262/K264/K266	Duodenal ulcer with bleeding and/or perforation
K275/K272/K274/K276	Peptic ulcer with bleeding and/or perforation
K285/K282/K284/K286	Peptic ulcer with bleeding and/or perforation
K290	Acute haemorrhagic gastritis
K625	Haemorrhage of anus and rectum
I850	Oesophageal varices with bleeding
I983	Oesophageal varices with bleeding in diseases classified elsewhere

Intracranial:

I60x	Subarachnoid haemorrhage
I61x	Intracerebral haemorrhage
I62x	Other non-traumatic intracranial haemorrhage

Other:

N02	Recurrent and persistent haematuria
H431	Vitreous hemorrhage
K661	Haemoperitoneum
R042	Haemoptysis
R040	Epistaxis
K250	Haemoarthrosis
N939	Vaginal bleeding

The ISTH (12) definition for major bleeding in non-surgical patients is defined by one of the following criteria:

1. Fatal bleeding, and/or

2. Symptomatic bleeding in a critical area or organ, such as intracranial, intraspinal, intraocular, retroperitoneal, intraarticular or pericardial, or intramuscular with compartment syndrome (from ICD10 codes in NPR), and/or

3. Bleeding causing a fall in hemoglobin level of 20 g/ L (1.24 mmol/L) or more, or leading to transfusion of two or more units of whole blood or red cells (from EMR data).

Organ system divided analyses (intracranial hemorrhage, gastrointestinal bleeding etc.) will be analyzed separately.

9.4 Data sources

For cohort 1 three data sources will form the basis for the analyses, the Norwegian Census Register (FRG), the Norwegian Prescription Database (NorPD) and the Norwegian Patient Register (NPR). The FRG list all Norwegian residents from birth to death, including migration data and place of birth.

The NorPD is a register covering all prescriptions dispensed at pharmacies nationwide and data is available from 1 January 2004. Each medication is coded according to the Anatomical Therapeutic Chemical (ATC) system. The prescription registry also includes information about date of dispensation, quantity dispensed, the defined daily dose (DDD) of the dispensed drug and strength. NorPD does not cover drug administered during hospital stays, drugs used at nursing home or over-the-counter drugs. From March 2007 the NorPD includes the reimbursement code using either the ICD10 or the ICPC-2 coding systems.

The NPR was established in 2008 and contains all hospital visits (emergency visits, in-patient hospitalizations and outpatient ambulatory consultations), length of stay and procedures (surgical and medical) from all hospitals in Norway. Diagnoses are coded according to the International Classification of Diseases, 10th revision (ICD10). Medical and surgical procedures are coded based on the NOMESCO (The Nordic Medico-Statistical Committee) coding system. Both primary and all subsequent codes related to each admission will be retrieved. Lag time in NPR is approximately 4 months.

The registries are nation-wide and provision of the information is mandatory, which eliminates the risk of both selection and recall bias. The large and detailed dataset also makes it possible to adjust for risk factors and analyze in depth the different sources of bias and confounding. The registry holders will generate the datasets and release it in a coded and de-identified form, but with a unique identifier common to the three datasets making individual merging of the NorPD, NPR and FRG datasets possible.

Data from the registries will be linked using the unique 11-digit national identification number. NorPD is a pseudonymized register and the registry holder will according to national guidelines handle the linkage and release the dataset in a coded form so that all participants remain de-identified.

The following key variables will be retrieved from FRG;

- Age
- Gender (male, female)
- Place of birth (Norway, Rest of the Nordic, Europe, Australasia, North America, South America)
- Emigration and immigration (country level)
- Death

The following key variables will be retrieved from NPR;

- Diagnoses codes: ICD10 codes
- Procedure codes: NOMESCO codes. NCMP for medical procedures and NCSP for surgical procedures
- Hospital stay details: date for admission, date of discharge, outpatient, in-patient stay, length of stay
- Type of hospital: University hospital, regional or local hospital

The following key variables will be retrieved from NorPD;

- Drug details: ATC codes, Article number, Name, Pack-size, Strength, Formulation and Defined
- Daily Dose (DDD) for all drugs.
- Dispensation details: date, number of packs, DDD for dispensation, category (reimbursed, non-reimbursed, institution etc.) and reimbursement codes. From 2008 prescribers in Norway need to write the reimbursement codes, using the conventional ICD10 or ICPC systems, on the prescription. Thus, we can from the reimbursement codes in NorPD exclude use of NOACs for other diagnoses such as atrial fibrillation

There is no official nation-wide mandatory registry for the primary care sector. However, dispensation of a prescription drug could be used as a proxy for an underlying diagnosis, e.g. dispensation of a diabetes type 2 medication indicates presence of a diabetes type 2 diagnosis. The ICD10/ICPC specific reimbursement code in NorPD also helps differentiating the use of a particular drug for several underlying diagnoses.

In cohort 2, patients will be identified through extraction of patient-level data from EMRs from a number of mid-sized to major hospital in Norway using the Pygargus Customized Extraction Platform (CXP) from IQVIA. This EMR extraction program has been used in a number of research projects since 2005 in the Swedish and Norwegian health care systems.(6, 7)

A key-code linked to each individual patient ID number will be extracted to enable linkage to national healthcare registers. This data will then be linked with the same national register data used in cohort 1. This key-code will be stored separate from the database used for analysis. Individual patient data from EMRs will then be linked to the described national registers (Norwegian Patient Registry, Norwegian Prescription Database). Extraction of patient level data will take place once and cover the period from 2009 to 30 June 2018 (or later depending on when EMR extraction can take place).

The following key variables will be retrieved from de-identified EMR data;

- Anthropometric data (height and weight)
- Laboratory values

- Deeper level longitudinal data on diagnoses and treatments, including referrals from primary care physicians.

These data will complement the data from the national registers.

9.5 Study size

There is no underlying hypothesis to confirm or reject. The sample is determined by eligibility. However it is desirable to assess the suitability of the available sample to produce useful estimates of the main outcomes.

The ratio of dispensations between the different NOACs and warfarin is estimated based on average number of dispensations in the period of 2014-2016 as given by NorPD statistics. Approximately half of patients treated with OAC have a dispensation for warfarin, 17%, 19% and 11% for apixaban, rivaroxaban and dabigatran respectively (Table 6). Based on previous studies, we estimate the risk of ischemic stroke and intracranial hemorrhage in the warfarin group, during an assumed mean follow-up of 18 months, to be 2.43% and 1.23% respectively.(1-4, 13-15)

ATC code and drug	**2013**	**2014**	**2015**	**2016**
B01AA03- warfarin	87,994	77,755	69,289	60,357
B01AF02 – apixaban	2,260	8,640	21,507	37,276
B01AF01 – rivaroxaban	13,423	20,795	25,480	28,924
B01AE07 - dabigatran	13,879	15,358	13,843	13,327
Total	117,556	122,548	130,119	139,884

Table 6: Number of patients with at least one dispensation of an OAC annually in the period 2013 to 2016 as given by NorPD. The table gives the total OAC use for all indications.

The study population will comprise all adult OAC naïve NVAF patients in Norway who filled a prescription for an OAC (rivaroxaban, apixaban, dabigatran, warfarin) in the study period, defined as from 1 January 2014 to 30 June 2018 (or later depending on availability of data). Based on Table 6 demonstrating the total number of new prescriptions for OACs in 2013-2016 and based on a publication describing Norwegian NVAF patients, the sample size of naïve patients available for the study will be approximately 70,000 patients; however, as known from previous unpublished data, the proportion of naïve patients who get prescribed a NOACs is likely to be increasing in the last several years.

The tables below show the precision for the proportions and relative risks of ischemic stroke and intracranial hemorrhage that might be expected in terms of width of 95% confidence intervals.

Assuming that 90% of dabigatran, apixaban or rivaroxaban patients can be propensity score matched with patients treated with warfarin, the 95% confidence intervals for the matched proportions show a sufficient precision for ischemic stroke and intracranial hemorrhage (see tables below).

	1:1 propensity score matched with warfarin		
Incidence rate	dabigatran N=6930	apixaban N=10710	rivaroxaban N=11970
1.46%	(1.19%, 1.77%)	(1.24%, 1.70%)	(1.25%, 1.69%)
1.94%	(1.62%, 2.29%)	(1.69%, 2.22%)	(1.70%, 2.20%)
2.43%	(2.08%, 2.81%)	(2.14%, 2.74%)	(2.16%, 2.72%)
2.92%	(2.53%, 3.34%)	(2.61%, 3.26%)	(2.63%, 3.24%)
4.37%	(3.90%, 4.88%)	(3.99%, 4.77%)	(4.01%, 4.75%)

Table 7: Two sided exact (Clopper-Pearson) 95% Confidence Intervals for the proportions of ischemic stroke; SAS 9.2;

The widths of the 95% confidence intervals for the incidence rates for ischemic stroke are not exceeding 0.98%, indicating a sufficient precision.

	1:1 propensity score matched with warfarin		
Relative Risk compared to warfarin	dabigatran N=6930	apixaban N=10710	rivaroxaban N=11970
0.60	(0.47, 0.77)	(0.49, 0.73)	(0.50, 0.72)
0.80	(0.64, 1.01)	(0.67, 0.96)	(0.68, 0.95)
1.00	(0.81, 1.24)	(0.84, 1.19)	(0.85, 1.17)
1.20	(0.98, 1.47)	(1.02, 1.41)	(1.03, 1.40)
1.80	(1.50, 2.17)	(1.55, 2.09)	(1.56, 2.07)

Table 8: 95% confidence intervals for relative risks for ischemic stroke with an assumed proportion of 2.43% for warfarin. The number of warfarin patients was assumed to be equal to the corresponding NOAC; SAS 9.2;

The widths of the 95% confidence intervals for the relative risks for the matched populations are not exceeding 0.67.

	1:1 propensity score matched with warfarin		
Incidence rate	dabigatran N=6930	apixaban N=10710	rivaroxaban N=11970
0.74%	(0.55%, 0.97%)	(0.58%, 0.92%)	(0.60%, 0.91%)
0.98%	(0.76%, 1.24%)	(0.80%, 1.19%)	(0.81%, 1.17%)
1.23%	(0.98%, 1.51%)	(1.03%, 1.46%)	(1.04%, 1.44%)
1.48%	(1.21%, 1.80%)	(1.26%, 1.73%)	(1.27%, 1.71%)
2.21%	(1.87%, 2.58%)	(1.94%, 2.51%)	(1.96%, 2.49%)

Table 9: Two sided exact (Clopper-Pearson) 95% Confidence Intervals for the proportions of intracranial hemorrhage; SAS 9.2;

The widths of the 95% confidence intervals for the incidence rates for intracranial hemorrhage are not exceeding 0.71%, indicating a sufficient precision.

	1:1 propensity score matched with warfarin		
Relative Risk compared to warfarin	dabigatran N=6930	apixaban N=10710	rivaroxaban N=11970
0.60	(0.42, 0.85)	(0.45, 0.79)	(0.46, 0.78)

0.80	(0.58, 1.10)	(0.62, 1.03)	(0.63, 1.02)
1.00	(0.74, 1.35)	(0.79, 1.27)	(0.80, 1.26)
1.20	(0.90, 1.60)	(0.95, 1.51)	(0.97, 1.50)
1.80	(1.38, 2.34)	(1.45, 2.22)	(1.48, 2.20)

Table 10: 95% confidence intervals for relative risks for intracranial hemorrhage with an assumed proportion of 1.23% for warfarin. The number of warfarin patients was assumed to be equal to the corresponding NOAC; SAS 9.2;

The widths of the 95% confidence intervals for the matched relative risks are not exceeding 0.96.

9.6 Data management

There will be 3 main steps;

- Merge together all data and build the analytical dataset
- Create patient cohorts
- Populate tables and figures

Statistical Analysis System (SAS software, version 9.4, SAS Institute, Cary, NC) will be used to all data handling to prepare analytical dataset.

The data from NorPD and NPR are considered high quality register data, however, some quality checking and data cleaning will be done to check for reasonability and consistency of data, e.g. by cross-tabulation of patient characteristics compared to publically available data sources and so forth. From these initial cut of data, further programming will be applied to identify the eligible patient cohorts with additional information of interest to address study objectives as include the following:

- All data from different sources will be retrieved/imported into SAS.
- Completeness of data regarding the number of variables and observations will be verified against the source.
- Quality control of the reasonability and consistency of information and data cleaning will be done
- Datasets will be merged to provide all information needed from various sources to define the patient cohorts.
- Patient cohorts will be defined and include all subjects which fulfills the cohort's inclusion criteria (section 9.2.2).
- Co-morbid conditions of interest within five years prior to index date will be identified, and various co-morbidity scores will be calculated (section 9.3.1).
- Concomitant medications at baseline and during the follow-up according to the ATC codes as described in section 9.3.1 will be selected.

- Adherence to each OAC will be estimated as described in section 9.3.3.
- All outcomes of interest (new events during follow-up) will be selected according to sections 8.1-3 and 9.3.4.1-2.
- Three different subset of cohorts for survival analysis based on exposure to OAC (continuous use or switched) will be defined as described in section 9.3.2.
- Finally a clean ready-to-use analysis database will be constructed, which will contain complete set of all data including the derived variables, according to the study protocol.
- Tables and figures will be populated

9.7 Data analysis

All data analyses will be performed using SAS statistical software version 9.4., and will be reported in a manner consistent with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.

Descriptive statistics will be used to present the data where continuous variables will be summarized as mean, standard deviation, median, minimum and maximum. Categorical variables will be described by frequencies and related percentages. Summaries will be reported separately for the elderly +75 years (cohort 1 and 2) and patients with renal impairment (cohort 2). Annual frequency of different OAC treatment during the study period in addition to adherence to each NOACs (high, moderate, low, and poor) will be presented. Also a separate analysis will be done to characterize patients with NVAF who have not been treated with OAC.

In the following, cohort 1 (registry) and cohort 2 (EMR) will be analyzed separately, but using the same analytical approaches.

The analyses will be performed on the total sample as well as on a confounder adjusted sample. It has to be kept in mind that analyses based on the overall sample (non-confounder adjusted sample) is prone to bias (see section 9.9).

Three separate propensity score models will be conducted and a 1:1 matching will be performed for:

- warfarin vs apixaban
- warfarin vs dabigatran
- warfarin vs rivaroxaban.

No matching to compare the NOACs among each other (e.g. apixaban vs rivaroxaban) will be performed due to underlying differences in prescribing preferences of different NOACs.

The propensity scores (PS) will be based on demographic variables and clinically relevant medical history, e.g. cardiovascular risk factors, comorbidities and concomitant medications at baseline.

Patients who switch or discontinue the OAC treatment will be censored at the date of switch or discontinuation. It cannot be ruled out that the switch of OAC (censoring) could be informative, but it is handled as uninformative censoring in the analyses.

Detailed methodology for propensity score modeling, matching, handling of missing data, summary and statistical analyses in this study will be documented in a Statistical Analysis Plan (SAP), which will be dated, filed and maintained by the sponsor. The SAP may modify the plans outlined in the protocol; any major modifications of primary endpoint definitions or their analyses would be reflected in a protocol amendment.

9.7.1 Analysis of primary outcome(s)

The incidence rates and corresponding 95% confidence intervals for ischemic stroke and intracranial hemorrhage will be calculated for warfarin, apixaban, dabigatran and rivaroxaban cohorts on a confounder adjusted population.

The person time at risk will be calculated from first OAC use during the study period until a certain outcome, death or end of follow-up in June 30, 2018 (or later depending on the availability of data). The incidence rates will be calculated by the number of first cases of a certain event during the follow-up divided by person time at risk and will be presented per 100 (or 1000) patient years with exact 95% Confidence Interval (Poisson distribution) for all outcomes of interest, and separately for each OAC.

To compare outcomes of interest in cohorts of warfarin with apixaban, dabigatran and rivaroxaban users, the incidence rate ratios and corresponding 95% confidence intervals will be calculated.

9.7.2 Further Analyses

The incidence rates and corresponding 95% confidence intervals for overall stroke (ischemic, hemorrhagic, other unspecified) and systemic embolism outcomes, myocardial infarction, all-cause mortality and major bleeding as defined by ISTH and the Cunningham algorithm (in a sub-set of patients; cohort 2) will be calculated for warfarin, apixaban, dabigatran and rivaroxaban cohorts.

Patients will be characterized according to demographic variables (age, gender), relevant medical history within 5 years prior to index date, different comorbidity risk scores (modified HAS-BLED, CHA2DS2VASC score, Charlson co-morbidity index), trends in OAC utilization from time of reimbursement, concomitant medications by OAC use and for NOACs-users, by the dose (standard or reduced) at start of follow up. Tests for differences between OAC sub-groups will be conducted using t-test, chi-square tests; the generalized linear model will be obtained to adjust for age and gender.

Additional analyses like a Cox proportional hazards regression models will be used to compare event rates of interest between the cohorts of users of different NOACs and warfarin adjusted for baseline covariates including age, gender, relevant medical history and concomitant medications. Hazard ratios with 95% confidence intervals for each study outcome will be presented (tables and Forest plots).

The occurrence of the first primary outcomes, as well as for the secondary outcomes among NVAF patients in warfarin versus NOACs will be described by Kaplan-Meier curves. The cumulative incidence of bleeding will be presented within different time windows (3m, 6m, 9m, 1 year, 2 years up until end of the follow-up).

Further details about the analysis and sensitivity analysis like Cox models evaluated within quintiles of PS probabilities can be found in the SAP.

9.7.3 Missing data

Missing data is not expected to be significant, as NorPD and NPD are both high quality national registers. But if there are any missing values, typically registered as NA (not applicable) in the data

files, these will be reported for all variables. If missing values should occur, and judged to be random, then imputation (if possible) will be carried out. Alternatively, missing values will be categorized as own category in multivariate analysis.

9.8 Quality control

The data from NorPD and NPR are considered high quality register data, however, some quality checking and data cleaning will be done to check for reasonability and consistency of data, e.g. by cross-tabulation of patient characteristics compared to publically available data and so forth. QC will be further detailed in the SAP.

The EMR data extracted in cohort 2 will undergo extensive cleaning and QC before analysis can take place.

9.9 Limitations of the research methods

The strength of the study is that it retrieves data from mandatory and nationwide registries in a public health care system that covers all residents. As a result, the dataset will give us a complete picture of all hospitalizations and prescriptions dispensed nationwide for the entire study period. This complete coverage of data eliminates also selection bias and recall bias that is an apparent problem using other databases being based on selected hospitals, health insurance schemes, or self-reported questionnaires. On top of the nationwide registry data EMR data will give us deeper level phenotypic data such as laboratory values (INR, Hb), weight/height, smoking and alcohol abuse.

This study will measure drug exposure at the level of pharmacy dispensation and will have no information on patient's real OAC intake.

There is also a risk of misclassification related to coding errors of hospital admissions; however, for serious conditions like bleeding and stroke this unlikely a problem. Within the scope of the study a formal validation of the AF diagnosis in NPR will be performed against EMR.

Channeling bias (confounding by indication and severity) pertaining to newer drugs given to poor responding or more fragile patients compared to established drugs is a potential limitation. To mitigate this risk of bias we will use warfarin as reference category and also start the study period 1 January 2014, which will omit the first 6-12 months of NOACs use after reimbursement approval, i.e. when prescription patterns have not yet been established.

Estimates based on the overall population are prone to bias because the decision to give one treatment over the other depends on different baseline characteristics. Therefore any differences in the estimates for some treatment may be caused by the differences in the populations rather by the given treatment.

Due to the separate estimation of the propensity scores the population to compare warfarin vs apixaban may be very different from the population to compare warfarin vs rivaroxaban. Therefore no comparison between the NOACs among each other can be made.

Nonetheless, and although adjustment for a number of important and know baseline characteristics will be performed, some unmeasured and residual confounding related to prescribing behavior or patient characteristics cannot be fully account for.

9.10 Other aspects

Not applicable

10. Protection of human subjects

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value and rigor and follow generally accepted research practices.

The project will store and preserve the de-identified dataset in accordance with Norwegian data privacy and registry holder requirements.

11. Management and reporting of adverse events/adverse reactions

As per the EMA Guideline on Good Pharmacovigilance Practices (Module VI–Management and reporting of adverse reactions to medicinal products), for non-interventional study designs that are based on secondary use of data, individual reporting of adverse reactions is not required. Reports of adverse events/reactions will be summarized in the study report (European Medicines Agency 2012).

12. Plans for disseminating and communicating study results

A Study Report will be completed comprising a full statistical report and a slide deck.

A primary manuscript will be prepared, in accordance with Bayer publication policy, in collaboration between study team members. Abstracts to scientific meetings will be submitted to first proper scientific meetings.

The author group will discuss and decide on whether one or several manuscripts will be written based on the results generated. The same goes for abstracts submitted to scientific meetings.

Development and preparation of study reports and publications will be a joint cooperation. Prior to commencing any publication activities, written agreements between the parties or author, need to be signed.

In accordance with the Declaration of Helsinki, both authors and publishers have ethical obligations. In the publication of the results of the study, the authors are obliged to preserve the accuracy of the

results. Negative as well as positive results should be published or otherwise made publicly available. All parties involved are committed to ensure that the data are reported in a responsible and coherent manner.

Publications generated will follow the guidelines established by the International Committee of Medical Journal Editors (ICMJE) and published in its Uniform Requirements of Manuscripts Submitted to Biomedical Journals. Authorship for all publications should comply with the criteria defined by the ICMJE. These state that: "Each author should have participated sufficiently in the work to take public responsibility for the content."

Each manuscript, abstract, oral or poster presentations are subject to internal [redacted] Publication team review and approval.

13. References

1. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med. 2009;361(12):1139-51.

2. Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, et al. Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med. 2011;365(11):981-92.

3. Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med. 2011;365(10):883-91.

4. Halvorsen S, Ghanima W, Fride Tvete I, Hoxmark C, Falck P, Solli O, et al. A nationwide registry study to compare bleeding rates in patients with atrial fibrillation being prescribed oral anticoagulants. Eur Heart J Cardiovasc Pharmacother. 2017;3(1):28-36.

5. Martinell M, Stalhammar J, Hallqvist J. Automated data extraction--a feasible way to construct patient registers of primary care utilization. Ups J Med Sci. 2012;117(1):52-6.

6. Lind M, Matsson PO, Linder R, Svenningsson I, Jorgensen L, Ploug UJ, et al. Clinical Effectiveness of Liraglutide vs Sitagliptin on Glycemic Control and Body Weight in Patients with Type 2 Diabetes: A Retrospective Assessment in Sweden. Diabetes therapy : research, treatment and education of diabetes and related disorders. 2016;7(2):321-33.

7. Rockberg J, Jorgensen L, Taylor B, Sobocki P, Johansson G. Risk of mortality and recurrent cardiovascular events in patients with acute coronary syndromes on high intensity statin treatment. Preventive medicine reports. 2017;6:203-9.

8. Fleet JL, Dixon SN, Shariff SZ, Quinn RR, Nash DM, Harel Z, et al. Detecting chronic kidney disease in population-based administrative databases using an algorithm of hospital encounter and physician claim codes. BMC nephrology. 2013;14:81.

9. Nielsen PB, Skjoth F, Sogaard M, Kjaeldgaard JN, Lip GY, Larsen TB. Effectiveness and safety of reduced dose non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. Bmj. 2017;356:j510.

10. Weir MR, Haskell L, Berger JS, Ashton V, Laliberte F, Crivera C, et al. Evaluation of clinical outcomes among nonvalvular atrial fibrillation patients treated with rivaroxaban or warfarin, stratified by renal function. Clinical nephrology. 2017.

11. Tamayo S, Frank Peacock W, Patel M, Sicignano N, Hopf KP, Fields LE, et al. Characterizing Major Bleeding in Patients With Nonvalvular Atrial Fibrillation: A Pharmacovigilance Study of 27 467 Patients Taking Rivaroxaban. Clinical Cardiology. 2015;38(2):63-8.

12. Schulman S, Kearon C, the SOCOAOTS, Standardization Committee Of The International Society On T, Haemostasis. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. Journal of Thrombosis and Haemostasis. 2005;3(4):692-4.

13. Graham DJ, Reichman ME, Wernecke M, Zhang R, Southworth MR, Levenson M, et al. Cardiovascular, bleeding, and mortality risks in elderly Medicare patients treated with dabigatran or warfarin for nonvalvular atrial fibrillation. Circulation. 2015;131(2):157-64.

14. Larsen TB, Skjoth F, Nielsen PB, Kjaeldgaard JN, Lip GY. Comparative effectiveness and safety of non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. BMJ. 2016;353:i3189.

15. Noseworthy PA, Yao X, Shah ND, Gersh BJ. Comparative effectiveness and safety of non-vitamin K antagonist oral anticoagulants versus warfarin in patients with atrial fibrillation and valvular heart disease. Int J Cardiol. 2016;209:181-3.

Annex 1: List of stand-alone documents

None

Annex 2: ENCePP checklist for post-authorization safety study (PASS) protocols

ENCePP Checklist for Study Protocols (Revision 2, amended)

Adopted by the ENCePP Steering Group on 14/01/2013; Doc.Ref. EMA/540136/2009

Study title:

PROTECT-AF: A **P**ost-marketing **R**etrospective n**O**n-interventional study using na**T**ionwide registries and electronic medical records to investigate the real-life **E**ffectiveness and major bleeding **C**omplications of oral an**T**icoagulants in Norwegian non-valvular **A**trial **F**ibrillation patients.

Study reference number: Not yet registered

Section 1: Milestones	**Yes**	**No**	**N/A**	**Page Number(s)**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[2]	☒	☐	☐	12
1.1.2 End of data collection[3]	☒	☐	☐	12
1.1.3 Study progress report(s)	☐	☐	☒	
1.1.4 Interim progress report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS register	☒	☐	☐	12
1.1.6 Final report of study results.	☒	☐	☐	12

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Page Number(s)**
2.1 Does the formulation of the research question and objectives clearly explain:				
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	12-13

[2] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.

[3] Date from which the analytical dataset is completely available.

Section 2: Research question	**Yes**	**No**	**N/A**	**Page Number(s)**
2.1.2 The objective(s) of the study?	☒	☐	☐	13-14
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	20-21
2.1.4 Which formal hypothesis (-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Page Number(s)**
3.1 Is the study design described? (e.g. cohort, case-control, randomised controlled trial, new or alternative design)	☒	☐	☐	14-15
3.2 Does the protocol specify the primary and secondary (if applicable) endpoint(s) to be investigated?	☒	☐	☐	34-37
3.3 Does the protocol describe the measure(s) of effect? (e.g. relative risk, odds ratio, deaths per 1000 person-years, absolute risk, excess risk, incidence rate ratio, hazard ratio, number needed to harm (NNH) per year)	☒	☐	☐	43

Comments:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Page Number(s)**
4.1 Is the source population described?	☒	☐	☐	18-21
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period?	☒	☐	☐	17
4.2.2 Age and sex?	☒	☐	☐	18
4.2.3 Country of origin?	☒	☐	☐	15-16
4.2.4 Disease/indication?	☒	☐	☐	18
4.2.5 Co-morbidity?	☒	☐	☐	22-31
4.2.6 Seasonality?	☐	☐	☒	

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Page Number(s)**
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	20-21

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
5.1 Does the protocol describe how exposure is defined and measured? (e.g. operational details for defining and categorising exposure)	☒	☐	☐	32-34
5.2 Does the protocol discuss the validity of exposure measurement? (e.g. precision, accuracy, prospective ascertainment, exposure information recorded before the outcome occurred, use of validation sub-study)	☒	☐	☐	32-34
5.3 Is exposure classified according to time windows? (e.g. current user, former user, non-use)	☒	☐	☐	32-34
5.4 Is exposure classified based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the product?	☐	☒	☐	
5.5 Does the protocol specify whether a dose-dependent or duration-dependent response is measured?	☒	☐	☐	32-34

Comments:

Section 6: Endpoint definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
6.1 Does the protocol describe how the endpoints are defined and measured?	☒	☐	☐	34-36
6.2 Does the protocol discuss the validity of endpoint measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, prospective or retrospective ascertainment, use of validation sub-study)	☐	☒	☐	

Comments:

Section 7: Confounders and effect modifiers	**Yes**	**No**	**N/A**	**Page Number(s)**
7.1 Does the protocol address known confounders? (e.g. collection of data on known confounders, methods of controlling for known confounders)	☒	☐	☐	44-45
7.2 Does the protocol address known effect modifiers? (e.g. collection of data on known effect modifiers, anticipated direction of effect)	☒	☐	☐	44-45

Comments:

Section 8: Data sources	**Yes**	**No**	**N/A**	**Page Number(s)**
8.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
8.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview, etc.)	☒	☐	☐	37-39
8.1.2 Endpoints? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics, etc.)	☒	☐	☐	37-39
8.1.3 Covariates?	☒	☐	☐	22-31
8.2 Does the protocol describe the information available from the data source(s) on:				
8.2.1 Exposure? (e.g. date of dispensing, product quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	37-39
8.2.2 Endpoints? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	37-39
8.2.3 Covariates? (e.g. age, sex, clinical and product use history, co-morbidity, co-medications, life style, etc.)	☒	☐	☐	22-31;37-39
8.3 Is a coding system described for:	☒	☐	☐	37-39

Section 8: Data sources	**Yes**	**No**	**N/A**	**Page Number(s)**
8.3.1 Diseases? (e.g. International Classification of Diseases (ICD)-10)	☒	☐	☐	37-39
8.3.2 Endpoints? (e.g. Medical Dictionary for Regulatory Activities (MedDRA) for adverse events)	☒	☐	☐	37-39
8.3.3 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC)Classification System)				
8.4 Is the linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	20-21

Comments:

Section 9: Study size and power	**Yes**	**No**	**N/A**	**Page Number(s)**
9.1 Is sample size and/or statistical power calculated?	☒	☐	☐	39-41

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Page Number(s)**
10.1 Does the plan include measurement of excess risks?	☒	☐	☐	42-43
10.2 Is the choice of statistical techniques described?	☒	☐	☐	42-43
10.3 Are descriptive analyses included?	☒	☐	☐	42-43
10.4 Are stratified analyses included?	☒	☐	☐	42-43
10.5 Does the plan describe methods for adjusting for confounding?	☒	☐	☐	42-43
10.6 Does the plan describe methods addressing effect modification?	☒	☐	☐	42-43

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Page Number(s)**
11.1 Is information provided on the management of missing data?	☒	☐	☐	41,43
11.2 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	41, 43-44
11.3 Are methods of quality assurance described?	☒	☐	☐	43
11.4 Does the protocol describe possible quality issues related to the data source(s)?	☒	☐	☐	43
11.5 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Page Number(s)**
12.1 Does the protocol discuss:				
12.1.1 Selection biases?	☒	☐	☐	43-44
12.1.2 Information biases? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods)	☒	☐	☐	43-44
12.2 Does the protocol discuss study feasibility? (e.g. sample size, anticipated exposure, duration of follow-up in a cohort study, patient recruitment)	☐	☒	☐	
12.3 Does the protocol address other limitations?	☒	☐	☐	43-44

Comments:

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Page Number(s)**
13.1 Have requirements of Ethics Committee/Institutional Review Board approval been described?	☒	☐	☐	44

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Page Number(s)**
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	44
13.3 Have data protection requirements been described?	☒	☐	☐	44

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Page Number(s)**
14.1 Does the protocol include a section to document future amendments and deviations?	☒	☐	☐	12

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Page Number(s)**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	45-46
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	45-46

Comments:

Name of the main author of the protocol: _____[redacted]_____

Date: / /

Signature: _____

Annex 3: Additional information

Annex 4: Signature pages