Non-Interventional Study Protocol

Study Protocol Title	Comparative Effectiveness and Safety of Direct Oral Anticoagulants in Patients with Nonvalvular Atrial Fibrillation in the UK
Version Identifier	Number 1.0
Date of Last Version	22 December 2021
Study Type	Non-Interventional Retrospective Comparative Effectiveness Study
Research Question and Objectives	This study aims to evaluate the incidence of stroke and other outcomes in association with direct oral anticoagulants (DOACs) as compared to each other (i.e., direct comparisons) among patients with nonvalvular atrial fibrillation (AFib) in the UK. Individual DOACs of interest include apixaban, rivaroxaban, edoxaban, and dabigatran. The primary objective is to ● Estimate the incidence rates and evaluate the association of stroke (ischemic or hemorrhagic) for patients with nonvalvular AFib who initiated apixaban compared to rivaroxaban The secondary objectives are to ● Compare the incidence rates of stroke (ischemic or hemorrhagic) for patients with nonvalvular AFib who initiated: ○ apixaban compared to edoxaban, dabigatran, and DOACs class ○ rivaroxaban compared to edoxaban, dabigatran, and DOACs class ○ edoxaban compared to DOACs class ○ dabigatran compared to DOACs class ● Estimate the incidence rates and evaluate the association of all-cause mortality, myocardial infarction, transient ischemic attack, major bleeding events, and major adverse cardiovascular events (MACE) for patients with nonvalvular AFib who initiated individual DOACs compared to those who initiated other DOACs.
Country of Study	United Kingdom
Study Review Board	Nicolle Gatto, PhD, MPH Chief Science Officer Aetion, Inc. Pall Jonnson, PhD Programme Director - Data NICE
Study Implementation Team	Ayad Ali, PhD, SRPharmS Senior Principal Scientist Aetion, Inc. Amanda Patrick, MS Senior Principal Product Scientist Aetion, Inc. Ashley Jaksa, MPH Scientific Partnerships Lead Aetion, Inc. Seamus Kent, PhD Senior Advisor in Data and Analytics NICE Shaun Rowark, MSc Scientific Advisor in Data and Analytics NICE Stephen Duffield, PhD, MBChB Senior Analyst NICE

	Liza Gibbs, BS Scientist Aetion, Inc.

1. Table of Contents

2. List of Abbreviations

Abbreviation	Description
AEP	Aetion Evidence Platform
AFib	Atrial fibrillation
ATC	Anatomical Therapeutic Chemical
CHA_2DS_2 VASc	Congestive heart failure, Hypertension, Age ≥75, Diabetes, prior Stroke/transient ischemic attack, vascular disease
CI	Confidence interval
CPRD	Clinical Practice Research Datalink
DOACs	Direct oral anticoagulants
EU	European Union
GCP	Good clinical practices
GP	General practitioner
GPP	Good pharmacoepidemiology practices
HdPS	High dimensional propensity score
HR	Hazard ratio
HTA	Health technology assessment
ICD-10	International Statistical Classification of Diseases 10th edition
ICH	International conference on harmonization

ICSR	Individual case safety report
ITT	Intent to treat
MACE	Major adverse cardiovascular events
MI	Myocardial infarction
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NSAID	Non-steroidal anti-inflammatory drugs
ORBIT-AF	Outcomes Registry for Better Informed Treatment of Atrial Fibrillation
QOF	Quality outcomes framework
RCA	Research collaboration agreement
RWD	Real-world data
RWE	Real-world evidence
STaRT-RWE	Structured template and reporting tool for real world evidence
THIN®	The Health Improvement Network
TIA	Transient ischemic attack
UK	United Kingdom
VKA	Vitamin K antagonists

3. Amendments and Updates

Amendment or Update No.	Date	Section of Study Protocol	Amendment or Update	Reason

4. Milestones

Milestone	Planned Date
Start of Data Extraction	September 2021
End of Data Extraction	November 2021
Diagnostics for primary comparison	December 2021
Registration in the EU PAS Register	January 2022
Interim Report *	January 2022
Final Report of Study Results	February 2022

* The interim report will comprise descriptive and primary objective results for primary comparison.

5. Rationale and Background

Health Technology Assessment (HTA) agencies, like the National Institute for Health and Care Excellence (NICE), are turning toward coverage with evidence development and/or managed access programs to allow additional time for data on the effectiveness of products to mature or be collected. Real-world data (RWD) has emerged as a potential strategy to collect additional evidence on the effectiveness of products after launch. HTA bodies are in the process of adopting real-world evidence (RWE), evidence generated from RWD, into their decision-making processes and setting standards for its use.

As NICE develops best practices on real-world evidence (RWE) use, it is exploring RWD comparative effectiveness studies as a way to enhance its decision-making and address uncertainties in its assessments and guidelines.[1] As part of NICE's five-year strategic vision,[1] NICE is initiating RWE research projects to fill known evidence gaps, to inform when and how RWE can be used in its decision-making, and to determine when after drug launch sample size is sufficient for comparative effectiveness analysis and if the results are consistent over time. NICE is interested in a proof-of-concept RWE comparative effectiveness study to inform these best practices.

5.1. Aetion's Research Collaboration with NICE

Aetion entered a research collaboration agreement (RCA) with NICE in 2021 to evaluate how RWE studies can be used to fill evidence gaps and reduce uncertainties in NICE assessment and guideline development. Aetion is collaborating with NICE to identify research questions relevant to NICE's RWE standards workstreams. Using the comparative effectiveness study described below, Aetion and NICE are collaborating on identifying fit-for-purpose real-world data, developing the protocol, executing the study using the Aetion Evidence Platform® (2021), software for real-world data analysis (Aetion, Inc. https://www.aetion.com.), and piloting the Structured Template and Reporting Tool for Real World Evidence (STaRT-RWE)[2] for transparent reporting of study implementation.

5.2. Selection of Atrial Fibrillation as Proof of Concept Study

The safety and efficacy of direct oral anticoagulants (DOACs) compared to vitamin K antagonists (VKAs) in patients with atrial fibrillation (AFib) have been established in randomized clinical trials.[3] In Europe, the following DOACs are approved to prevent venous thromboembolism, deep vein thrombosis, pulmonary embolism, and stroke in adults with AFib: Apixaban (Eliquis®, first authorised in the EU and the UK May 2011); dabigatran (Pradaxa®, March 2008); rivaroxaban (Xarelto®, September 2008); and edoxaban (Lixiana®, June 2015).[4,5]

Between March 2012 and September 2015, NICE separately assessed and recommended four DOACs: dabigatran, rivaroxaban, apixaban, and edoxaban. Each submitted clinical trial compared the investigational DOACs to warfarin; however, there were no head-to-head clinical trials comparing the DOACs to each other at the time of NICE's assessments. Therefore, a network meta-analysis was completed for indirect comparison of individual DOACs agents using warfarin as the common comparator. NICE interpreted the network meta-analysis findings with caution noting the results were not sufficiently robust to differentiate between the products' effectiveness. These uncertainties were reiterated in NICE's 2014 AFib management guidelines in 2014 and 2021 updates.[6,7]

The lack of relevant direct comparisons driving uncertainty is common at the time of NICE assessment, and these uncertainties are not unique to AFib. The uncertainties in comparative effectiveness of AFib treatments impact downstream decisions by healthcare providers and patients as they try to identify the most effective, safe, and cost-effective treatment option for stroke prevention.

NICE has selected AFib for the proof-of-concept comparative effectiveness real-world effectiveness (RWE) study in part because there is a growing body of literature that has successfully validated the use of real-world data (RWD) in AFib,[8,9] and the burden of disease and cost to the healthcare system associated with AFib in the UK.[7] In addition, the DOACs for nonvalvular AFib included in this study were approved more than 5 years ago in the UK (edoxaban was approved in June 2015[10]), which allows us to explore how the value of RWE comparative effectiveness changes with length of follow-up. This can help inform optimal timing for reassessments.

This study will not only inform when and how NICE could use RWE to address uncertainties in the clinical evidence, but it could also inform NICE's AFib clinical guidelines.

5.3. DOACs Therapies for Atrial Fibrillation

Atrial fibrillation (AFib) is the most common type of cardiac conduction disorder accounting for up to 2.6% of healthcare costs in Europe, and about 9.5% of the European population is estimated to have AFib by 2060.[11] In the UK, AFib affects 1.4 million patients,[12] and between 0.9% to 1.6% of UK's National Health Service (NHS) spendings are attributable to AFib predominately from hospitalizations.[13] The condition is associated with significant complications, including stroke. Nonvalvular AFib increases an individual's risk of stroke by about five times,[14] and between 20% and 30% of stroke cases are attributed to AFib.[11]

Anticoagulants, including VKAs and DOACs, are highly effective in the prevention and treatment of thromboembolic events associated with AFib. Nevertheless, VKAs like warfarin are narrow therapeutic index drugs requiring frequent coagulation monitoring, and have multiple drug and food interactions. Alternatively, DOACs inhibit coagulation via direct and specific binding to active sites of thrombin (e.g., dabigatran) or factor Xa (e.g., apixaban, rivaroxaban, and edoxaban) of the coagulation pathway. Compared to VKAs, DOACs have a wider therapeutic index, which permits use in fixed doses without coagulation monitoring, and relatively limited drug and food interactions. Data from direct comparisons of DOACs (e.g., apixaban vs. dabigatran) are not available, and relative safety and efficacy findings are based on indirect comparisons from network meta-analyses.[15,16] Healthcare providers face challenges from uncertainties around DOACs treatment decisions for patients with AFib, particularly around which individual agent has the best benefit-risk profile within the class for mortality and cardiovascular risk reduction.

The current protocol details a study designed to evaluate the effectiveness and safety of the DOACs of interest (apixaban, rivaroxaban, edoxaban, and dabigatran) in direct comparisons, among adults with nonvalvular AFib at risk for stroke in the UK.

6. Research Objectives

This study evaluates the comparative effectiveness of apixaban versus rivaroxaban (primary objective) for reducing stroke among patients with nonvalvular AFib in the UK.

Secondary objectives included comparing the effectiveness of apixaban and rivaroxaban to the other DOACs (edoxaban, and dabigatran).

6.1. Hypothetical Target Trial

When planning and designing RWE studies, it is important to consider emulating a hypothetical target trial.[17] We sought to emulate a four-arm randomized trial that randomizes patients to apixaban, rivaroxaban, edoxaban, or dabigatran. The objective of the hypothetical trial is to compare apixaban to rivaroxaban, as the primary comparison, in the reduction of stroke among patients with nonvalvular AFib in the UK (Figure 6.1). The hypothetical target trial diagram is based on four clinical trials that evaluated the efficacy of respective DOACs in comparison with warfarin in patients with AFib (ARISTOTLE trial studied apixaban;[18] ROCKET AF studied rivaroxaban;[19] ENGAGE AF-TIMI 48 trial studied edoxaban;[20] and RE-LY trial studied dabigatran).[21] In addition, comparisons of individual DOACs to other DOACs as a pharmacological class will be included as secondary comparisons. The hypothetical target trial would be powered for the primary comparison of apixaban versus rivaroxaban, and the study designed to meet diagnostic criteria for the main analysis.

Within the UK, apixaban and rivaroxaban are the most frequently prescribed DOACs (Figure 6.2). Secondary comparisons will include comparing apixaban and rivaroxaban to the two other DOACs (edoxaban and dabigatran) (Table 6.1).

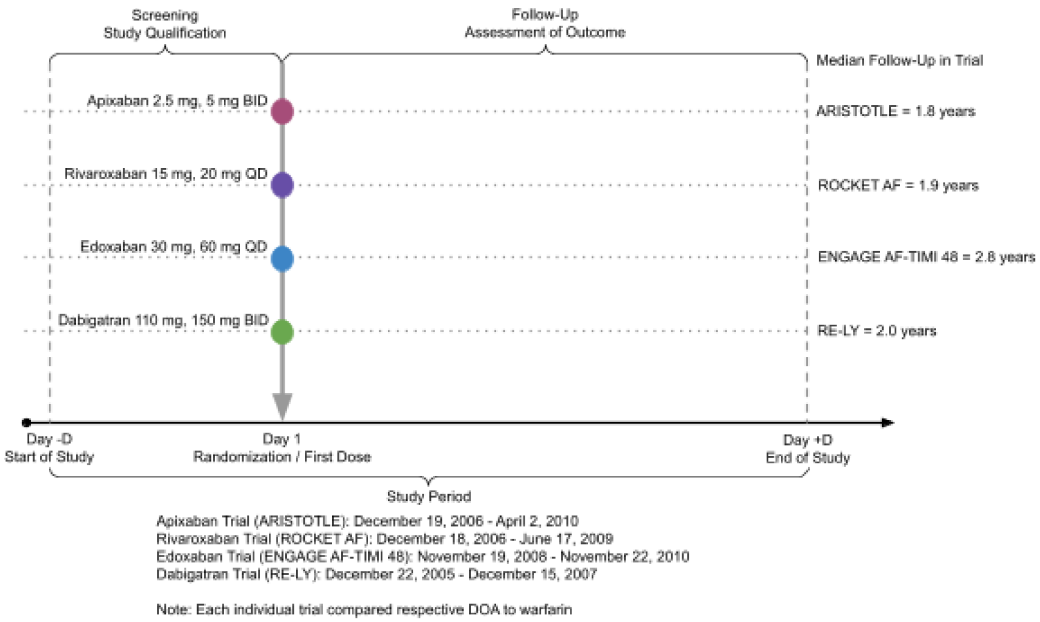

Figure 6.1 Hypothetical Target Trial Design

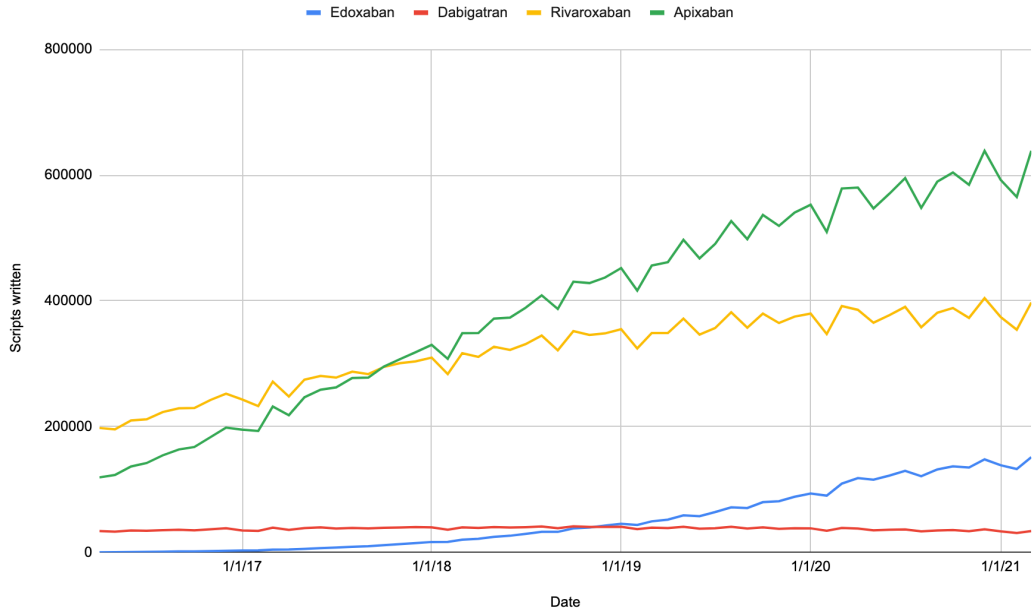

Figure 6.2 Utilization of DOACs in the UK (April 2016 - January 2021; Source: OpenPrescribing)

Table 6.1 Direct Comparisons of Study Exposures of Interest

	Rivaroxaban	Edoxaban	Dabigatran	DOACs Class
Apixaban	Primary	Secondary	Secondary	Secondary (rivaroxaban, edoxaban, dabigatran)
Rivaroxaban	-	Secondary	Secondary	Secondary (apixaban, edoxaban, dabigatran)
DOACs Class	-	Secondary (apixaban, rivaroxaban, dabigatran)	Secondary (apixaban, rivaroxaban, edoxaban)	-

6.2. Primary Objective

The primary objective is to estimate the incidence rates and analogous hazard ratio of stroke (ischemic or hemorrhagic) for patients with nonvalvular AFib who initiated apixaban compared to rivaroxaban (see Table 6.1).

6.3. Secondary Objectives

The secondary objectives are to:

- Estimate the incidence rates and analogous hazard ratios of stroke (ischemic or hemorrhagic) for patients with nonvalvular AFib who initiated:
    - apixaban compared to edoxaban, dabigatran, and DOACs class
    - rivaroxaban compared to edoxaban, dabigatran, and DOACs class
    - edoxaban compared to DOACs class
    - dabigatran compared to DOACs class

- Estimate the incidence rates and evaluate the analogous hazard ratio of all-cause mortality, myocardial infarction (MI), transient ischemic attack (TIA), major bleeding events, and major adverse cardiovascular events (MACE) for patients with nonvalvular AFib who initiated individual DOACs compared to those who initiated other DOACs (see Table 6.1).

7. Data Source

This study uses anonymized patient RWD from The Health Improvement Network (THIN®) Database (A Cegedim Proprietary Database). THIN® is an unobtrusive medical data collection scheme that collects anonymised patient data from its members. THIN® data is a primary care research database containing anonymised electronic health record data from around 850 UK general practices (GPs) using the VISION clinical system (since 1994), containing records around 20 million patients, of which 2.6 million are currently active. THIN® collates the computerized medical records of general practitioners (GPs) in the UK who act as the gatekeepers of healthcare and maintain patients' life-long electronic health records. As such they are responsible for primary healthcare and specialist referrals, and they also store information stemming from specialist referrals, and hospitalizations. GPs act as the first point of contact for any non-emergency health-related issues, which may then be managed within primary care and/or referred to secondary care as necessary. The THIN® database includes demographic, medical diagnosis, patient medical history, lifestyle factors, and written prescription information on individual patients. The Quality and Outcome Framework (QOF) rewards practices for the quality care they provide their patients, based on a number of indicators, including medical record data entry.

The THIN® database has been used for scientific publications in AF and stroke research[22-25] and has been validated and widely used in pharmacoepidemiologic studies.[22]

8. Research Methods

8.1. Study Design

This study is a non-interventional, retrospective cohort study utilising RWD from the UK to address the study objectives. Patients with nonvalvular AFib who initiated (i.e., new users of) DOACs during the observation period will comprise the study cohort. The exposure groups of interest will be patients initiating individual DOACs (apixaban, rivaroxaban, edoxaban, and dabigatran). All patients will be followed for the first occurrence of each of the study outcomes of interest.

8.2. Setting

This study will be conducted using an electronic anonymized patient RWD from The Health Improvement Network (THIN®) Database (A Cegedim Proprietary Database) that includes data collected in primary care settings in the UK.

8.2.1. Study Population

This study will be conducted using patient clinical data extracted from the THIN® database in the UK for adults with nonvalvular AFib at risk for stroke who initiated DOACs between July 1, 2014 and December 31, 2020. This period was selected to minimize potential effects on recommended treatment options due to changes in AFib management guidelines. The NICE AFib management guidelines, which recommend prescribing a DOACs (apixaban, rivaroxaban, edoxaban, or dabigatran) for stroke prevention were published in June 2014[6] and were not updated until April 2021[7], thus we believe treatment options to be stable during this period. Patients with AFib are identified by Read Medical Codes and ICD-10 diagnosis codes.

Patient selection criteria are described in the following sections and Figure 8.1.

8.2.1.1. Inclusion Criteria

The following are the inclusion criteria for the study cohort:

- ≥1 prescription for DOACs of interest (index date)

8.2.1.2. Exclusion Criteria

The following are the exclusion criteria for the study cohort:

- Exclusion criteria applied during selection of the data cut:
    - <1 diagnosis code for AFib on index date or in 12 months prior (applied during selection of the data cut)
    - <1 risk factor for stroke other than Afib on index date or in 12 months prior defined by CHA_2DS_2 VASc >1 for men and >2 for women[26]
    - <1 medical encounter in the 180 days prior to index date
    - ≥1 diagnosis code for the following conditions on index date or any time prior:
        - Cardiac valve disease (mitral stenosis, valvular replacement)
        - Deep vein thrombosis
        - Pulmonary embolism
- Exclusion criteria applied after selection of the data cut:
    - <18 years of age on index date
    - <12 months' registration with a GP prior to the index date
    - ≥1 prescription record for DOACs in the 12 months prior to the index date (i.e., prevalent users)
    - ≥1 diagnosis code for the study outcomes of interest (stroke, MI, TIA, major bleeding events) on index date or any time prior
    - ≥1 diagnosis code for the following conditions on index date or any time prior:
        - Angina
        - Congenital heart disease
    - Missing age
    - Missing gender

8.2.2. Study Periods

The study observation period spans July 1, 2013 through December 31, 2020, consisting of a drug indexing, baseline and follow-up periods (Figure 8.1). The drug indexing period will be used to ascertain DOACs initiation (index date). Baseline period is defined as 12 months prior to and including the index date. The follow-up period will include all available time after study index date. Patients will be followed in an “intention to treat” (ITT) approach from their index date until the end of follow-up period or the occurrence of the following events, whichever occurs first:

- First occurrence of the study outcomes of interest
- Death
- End of patient data
- End of patient registration with GP.

In addition, patients meeting the selection criteria will be followed in an “as-treated” approach from their index date until the end of the study follow-up period or the occurrence of the following events, whichever occurs first:

- First occurrence of the study outcomes of interest.
- Death
- End of patient data.
- End of patient registration with GP.
- Termination of exposure, crossover of exposure group, or addition of another DOAC.

Termination of exposure in the as-treated analysis is defined as having a gap exceeding 30 days (grace period) between the end of a prescription, based on its start date and duration, and the start of the next prescription, or if no additional prescription occurs. The termination date is defined as the prescription end date plus a 30-day risk window.

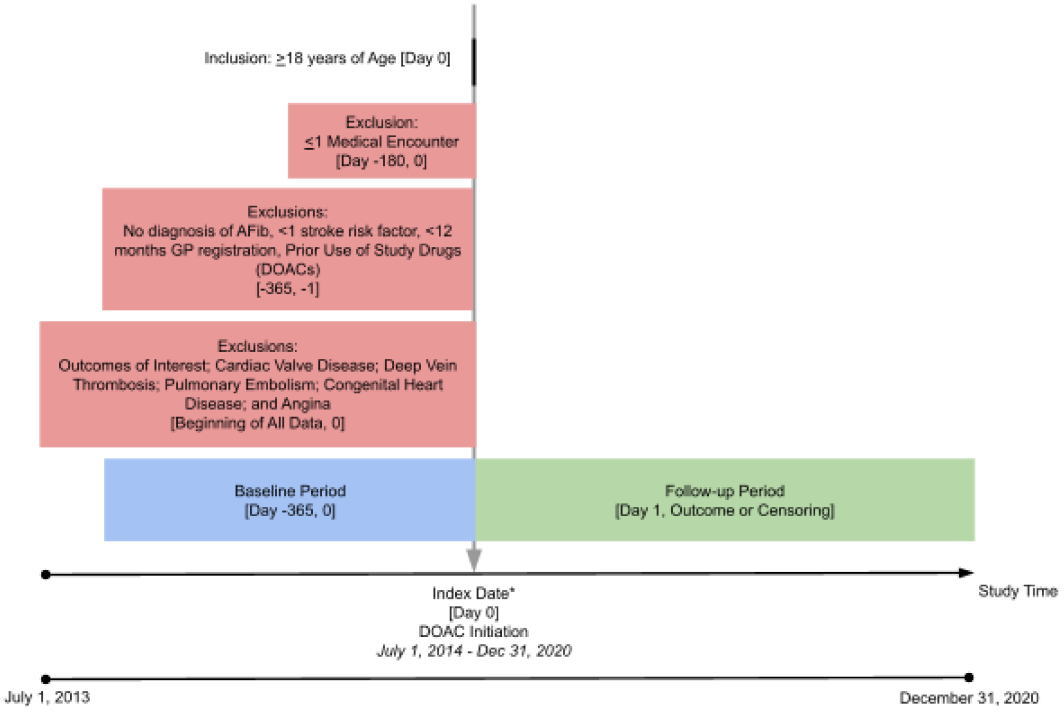

Figure 8.1 Study Design Diagram

8.3. Variables

The following sections describe the study exposures, outcomes, and covariates (including patient demographics, clinical characteristics, comorbidities, and concomitant medications).

8.3.1. Exposures

DOACs of interest will include apixaban, rivaroxaban, edoxaban, and dabigatran. Exposures will be defined by corresponding ATC codes: B01AF02; B01AF01; B01AF03; and B01AE07, respectively. Initiators of DOACs are defined as new users of DOACs with no record of prescription for any DOAC during 12 months before index date.

Only a single index date is allowed per patient; if a patient qualifies as a new user of an exposure group multiple times, the patient will enter the cohort on the first qualifying exposure date. Patients with exposure to more than one DOAC on the index date will be excluded. Duration of treatment is calculated from the provided days supplied/duration data field. Fields with 0 are assumed to be 1 day. A grace period of 30 days between refills and risk window of 30 days after the last refill is used to define persistence.The operationalized definition of exposure will be recorded in the STaRT RWE template and is available on request.

8.3.1.1. Primary Exposure Groups

- *Exposure Group 1:* Initiators of apixaban
- *Exposure Group 2:* Initiators of rivaroxaban

8.3.1.2. Secondary Exposure Groups

The following independent Exposure Groups will be created:

- *Exposure Group 3:* Initiators of edoxaban
- *Exposure Group 4:* Initiators of dabigatran
- *Exposure Group 5:* Initiators of rivaroxaban, edoxaban, and dabigatran
- *Exposure Group 6:* Initiators of apixaban, edoxaban, and dabigatran
- *Exposure Group 7:* Initiators of apixaban, rivaroxaban, and dabigatran
- *Exposure Group 8:* Initiators of apixaban, rivaroxaban, and edoxaban

8.3.1.3. Primary Comparison

- *Primary Comparison:* initiators of apixaban vs initiators of rivaroxaban (i.e., *Exposure Group 1* vs *Exposure Group 2*)

8.3.1.4. Secondary Comparisons

- *Secondary Comparison 1:* initiators of apixaban vs initiators of edoxaban (i.e., *Exposure Group 1* vs *Exposure Group 3*)
- *Secondary Comparison 2:* initiators of apixaban vs initiators of dabigatran (i.e., *Exposure Group 1* vs *Exposure Group 4*)
- *Secondary Comparison 3:* initiators of apixaban vs initiators of other DOACs class (i.e., *Exposure Group 1* vs *Exposure Group 5*)
- *Secondary Comparison 4:* initiators of rivaroxaban vs initiators of edoxaban (i.e., *Exposure Group 2* vs *Exposure Group 3*)
- *Secondary Comparison 5:* initiators of rivaroxaban vs initiators of dabigatran (i.e., *Exposure Group 2* vs *Exposure Group 4*)
- *Secondary Comparison 6:* initiators of rivaroxaban vs initiators of other DOACs class (i.e., *Exposure Group 2* vs *Exposure Group 6*)
-
- *Secondary Comparison 7:* Initiators of edoxaban vs initiators of other DOACs class (i.e., *Exposure Group 3* vs *Exposure Group 7*)
- *Secondary Comparison 8:* initiators of dabigatran vs initiators of other DOACs class (i.e., *Exposure Group 4* vs *Exposure Group 8*)

8.3.2. Outcomes

The primary outcome is ischemic or hemorrhagic stroke. Secondary outcomes include the occurrence of all-cause mortality, myocardial infarction, transient ischemic attacks, major bleeding events, and major adverse cardiovascular events (MACE). Major bleeding is defined as a composite outcome of major intracranial (including hemorrhagic stroke), gastrointestinal, and urogenital bleeds. MACE outcomes will be a composite endpoint of angina, myocardial infarction, and stroke.

Outcomes of interest will be measured during the study follow-up period, and defined by corresponding Read Medical Codes and ICD-10 diagnosis codes at the primary care setting. The operational definitions of each outcome will be recorded in the STaRT-RWE template and are available on request.

8.3.3. Covariates

Patient demographics, clinical characteristics, comorbidities, and concomitant medications will be measured during baseline period and index date, and identified by Read Medical Codes,

ICD-10 diagnosis codes, and medication ATC codes for corresponding covariates. These covariates include cohort entry year, age, sex, socioeconomic indicators (e.g., marital status), and selected components of CHA_2DS_2 VASc stroke risk score[27,28] (heart failure, hypertension, and diabetes), and ORBIT-AF bleeding risk score in AFib patients[29,30] (heart failure, renal impairment, hepatic impairment; osteoporosis/hip fractures, anemia, antiplatelet therapy, aspirin therapy, NSAIDs therapy, and smoking and alcohol drinking history [if available]).

Additional comorbidities and concomitant medications deemed to be associated with the primary and secondary outcomes of interest will be included in the analyses. All covariates will be determined based on the literature and clinical knowledge. The operational definitions of each covariate will be recorded in the STaRT-RWE template and are available on request.

9. Data Analysis

Data analyses will be performed in a stepwise approach involving three distinct sequential phases (Figure 9.1): *Exploratory Phase* (data explorations will be done in advance to inform key design decisions), *Diagnostic Phase* (requirements that must be met prior to viewing study outcomes, e.g. covariate balance), and *Inferential Phase* (comparative analyses).

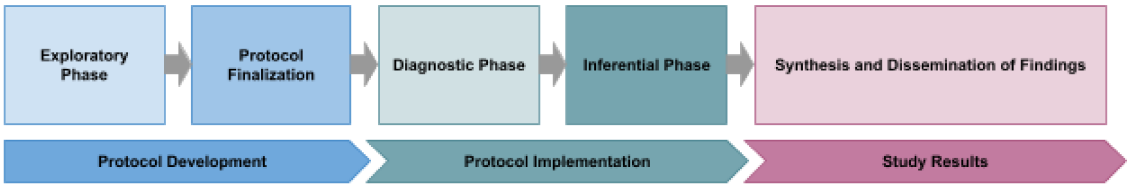

Figure 9.1 Data Analysis Phases

In the *Exploratory Phase*, the study implementation team completed explorations related to the patient selection criteria, exposure groups, and outcome definitions that informed the development of the study protocol.

For the *Diagnostic Phase*, the checklist in Table 9.1 must be satisfactorily completed prior to beginning the implementation of the *Inferential Phase*. As such, the relationship between the exposure and outcome of interest will not be described or evaluated in the analytic dataset until the study implementation team reaches consensus that the diagnostic criteria are satisfied.

Table 9.1 Diagnostic Checklist to Complete Prior to Progression to the Inferential Phase

Diagnostic Criteria	Description
Confirm adequate sample size	Adequate sample size will be defined as the minimum sample size required to achieve 80% power to detect a HR of ≥0.80 for the primary outcome of stroke under the assumptions of 1:1 matching ratio of the primary comparators (apixaban vs. rivaroxaban); a false-positive rate α=0.05; a 1-sided test; and a background stroke incidence rate of 47 events per 100,000 person-years in patients with AFib[31] Note: If adequate sample size is not reached (i.e., minimum HR

	that we can detect is <0.80), it will not solely prohibit us from moving on to interferential analysis. NICE would typically use treatment effect estimates to parameterise economic models and as long as the uncertainty in estimates is quantified, it can be accounted for in decision-making. In addition, a recent article[32] deprioritized the importance of power calculations for observational studies, noteing that observational analyses with imprecise effect estimates should be completed as small studies can be pooled to provide a more precise effect estimate. Therefore, this analysis is informative even if the sample size is insufficient to detect a HR ≥0.80 with 80% power.
Confirm positivity of variables	Propensity score distributions will be visually inspected, and overlap in all areas of the propensity score distributions will be confirmed. Other approaches, such as propensity score weighting, will be explored to achieve covariate balance in case propensity score matching is insufficient.
Confirm baseline confounder balance	The distributions of all potential confounders will be confirmed to be balanced for each comparison of interest. Covariate balance will be defined as ASD ≤0.10. Although covariates with balance prior to propensity score matching may be removed from the propensity score model, balance of these covariates will still be confirmed after matching. Covariates with small residual imbalance (defined as 0.10 ≤ ASD ≤ 0.15) may be deemed balanced if the covariate does not predict the outcome among the referent group (defined as ASD <0.10 when comparing the risk of the outcome in those with the covariate vs. those without it) Evaluation of imbalance in outcome prediction will only be conducted once at the end of the diagnostic phase if all other diagnostic criteria are met.
Confirm models are not overfit	All models must contain ≥12 exposed patients per covariate
Confirm persistence of treatment	Initiators of DOACs stay on treatment for a substantial amount of time after starting treatment. The definition of “substantial persistence” will be finalized during the diagnostic phase and will be data driven. If substantial crossover or censoring is observed, appropriate methods to account for these censoring issues will be applied in the comparative analyses. Preliminary analyses show patients treated with DOACs in the database have a median follow-up of 2.3 years (IQR, 1.0-3.9). In clinical trials, the median duration of treatment was 1.5 years.

Unless otherwise specified, results will be provided as descriptive statistics with categorical variables reported using frequency distributions, and continuous variables reported using means, standard deviations, medians, minimums, maximums, 25th percentiles and 75th

percentiles. For inferential statistics, both crude and covariate adjusted analyses will be reported for study outcomes.

Among patients in each of the exposure groups, the incidence rates and 95% confidence intervals (CI) of stroke and secondary outcomes will be calculated for each of the mutually exclusive comparison groups, along with the number of events, total number of individuals, and accrued person-time. Incidence rates will be calculated as the number of incident outcomes of interest within the follow-up period divided by the total person-time at-risk, and reported as rate per 1,000 person-years.

In each comparison cohort, multivariable logistic regression will be used to estimate the probability of a patient's initiation of the exposure given baseline covariates and cohort entry year (i.e., the propensity score). Propensity score matching between exposure groups will be performed using 1:1 nearest neighbor matching without replacement with a maximum matching caliper of 0.01. In addition to graphical depictions of propensity score distributions, the absolute standardized differences (ASD) in proportions and means of baseline characteristics will be estimated to examine comparability of exposure groups. Covariates with ASD ≤0.10 are considered balanced between comparator groups.[33] Covariates with small imbalances 0.10 ≤ ASD ≤ 0.15) may be deemed balanced if the covariate does not predict the outcome among the referent group.

Cox proportional hazards regression (outcomes model) will be used to estimate hazard ratios (HR) and 95% CI for each outcome after propensity score matching. The incidence of stroke and secondary outcomes will be compared between individual DOACs in primary and secondary comparisons as mutually exclusive cohorts.

Subgroup analyses that will be considered include DOACs initiators with prior warfarin use; age (<75 and ≥75 years); concomitant aspirin use; CHA_2DS_2 VASc score (0-1, 2-3, and ≥4), and selected comorbid conditions, e.g. diabetes, heart failure, and BMI status (<30 and ≥30 kg/m²).

Sensitivity analyses will be performed to assess the impact of varying the study period on study findings by limiting analyses on data up to February 28, 2020 to account for potential impact of COVID-19 pandemic on healthcare utilization.

Additionally, high-dimensional propensity score (HdPS) analysis will be used as a sensitivity analysis to estimate the association between treatment with DOACs and the primary outcome of stroke. The HdPS approach is a 7-step algorithm that empirically identifies a pool of covariates from different data dimensions based on their prevalence, and then selects a subset of the covariates for inclusion in a propensity score model based on their potential to bias the exposure-outcome association.[30]

Missing values will be reported as missing, and no imputation will be undertaken. All data analysis will be conducted using Aetion Evidence Platform® (2021), software for real-world data analysis. Aetion, Inc. https://www.aetion.com.

10. Limitations of the Research Methods

This study will be based on secondary use data from EHR with inherent limitations, including misclassification. For example, a patient may not necessarily use a medication that they have received a prescription for; however, this is less likely in the case of medications for chronic conditions like cardiovascular diseases. Additionally, while it is possible to identify study

outcomes of interest in the THIN® database, misclassification of the outcome is possible, as the presence of a diagnosis in medical records does not necessarily indicate a true presence of disease; however, the outcomes of interest are included in the Quality and Outcome Framework (QOF)[34] which rewards and incentivizes GPs in England for quality care and coding of diagnoses. In addition, validated algorithms from previously published real-world studies will be used to define study variables of interest.

An incident user design will be applied to reduce biases that can impact non-randomized studies, especially when using EHR.[35] An ITT analysis will be used because it has the advantage of eliminating certain types of biases by preserving the prognostic balance obtained through the propensity score matching and maintaining sample size;[36,37] however, ITT analysis is sometimes considered to be “too conservative” and has the possibility of exposure misclassification.[35,36] Additionally, “as-treated” approach will be used that will censor individuals upon discontinuation of index treatment or treatment crossover. However, limitations of an “as-treated” analysis include the possibility of introducing certain biases including differential and informative censoring, or time-dependent confounding biases.[36,37]

Finally, while the THIN® database contains substantial information for inclusion/exclusion criteria and confounder control, medical conditions or a family history of medical conditions are only ascertainable where established diagnoses and procedures for those conditions exist. Additionally, some key covariates may not be available (unmeasured or imperfectly measured confounders), e.g. alcohol and smoking histories are not routinely recorded. Residual confounding by unmeasured variables will likely be present; however, efforts will be made to incorporate all potential confounding variables into the propensity score analysis to minimize impacts from unmeasured confounding, including the application of HdPS sensitivity analysis. Compared to conventional confounding adjustment methods, HdPS algorithm improves confounding control in situations when the variables are weak confounders that are weakly associated with exposure and themselves are associated with unmeasured confounders.[38]

11. Quality Control

The study data to be used for the proposed analysis will be connected to the Aetion Evidence Platform (AEP), which will be used to perform all the analyses. The AEP is a data-handling technology, which allows for the analysis of large patient claims, EHR, and other transactional datasets by indexing patient data into a form that can be queried by an internal patient variable language. Data is minimally transformed at the point of connection to the AEP, thus the original format of the THIN® data is preserved. At the point of data connection to the platform some discard rules are applied. Patient events are excluded if there are no dates associated with them, or if the start date of the event is preceded by the end date of the event (e.g. discharge date precedes admission date for an inpatient event). Aetion IDs are assigned to THIN® patient IDs and a crosswalk file is kept as a protected file available upon request to authorized parties. The patient data is individual level patient data and will be analyzed within the AEP. Aggregated results will be exported from the platform in the form of tables.

THIN® data are loaded into the AEP after minimal processing into patient longitudinal timelines to enable representation of the original data and without any data loss. The following data checks are performed during the data connection process:

- Events are required to have a valid start date.
- Record counts are cross-checked for validation and compared to the original data counts.

- Enrollment information is processed and checked, and all fields and sample records are checked to ensure validity.

The process is implemented via a double-programming, rule-based approach that is flexible, automated, and scalable so that it can be reproduced when data is updated. Data processed through the data connector (coded by the Aetion Data Engineering team) are compared against data transformed independently to reproduce the data connector logic (coded by the Aetion Data Science team). Checks are performed first on a subset of data, then on the full data. This process ensures that customers are working with a scientifically valid data product when they perform analyses on the data using the AEP. Statistical analyses will be conducted using the validated AEP.

The Aetion Science Team will build measures for cohort inclusion/exclusion criteria, outcomes and covariates. All measures created, cohorts developed, statistical analyses implemented, and tables populated will undergo quality control review by at least one additional analyst or scientist under the supervision of the Senior Scientist. Quality control methods include checks for the validity and logical content of codes and checks for missing values and variables. In order to control for potential inconsistencies and errors, all variables will be tabulated. In addition, the distribution of values for each variable, including potential outliers, will be examined. This protocol will be strictly followed in the study implementation. However, variable definitions may undergo modification if determined to be scientifically sensible. All changes to this protocol will be documented in protocol amendments.

12. Protection of Human Subjects

This study will be conducted in accordance with applicable laws and regulations of the country where the study is being conducted, as appropriate. This observational, non-interventional study does not affect the treatment of the patients. The study is conducted in accordance with ethical principles that are consistent with the Declaration of Helsinki, International Conference on Harmonisation (ICH) Good Clinical Practice (GCP), Good Pharmacoepidemiology Practice (GPP), and the applicable legislation on non-interventional studies and/or observational studies.

The study concept was approved by the THIN® Scientific Review Committee on July 6, 2021.

13. Management and Reporting of Adverse Events/Adverse Reactions

This is a non-interventional study using only structured secondary data, and attribution of adverse reactions/adverse events to specific exposure is not possible, and therefore no individual case safety report (ICSR) reporting is required. During the course of observational research using existing secondary databases, the proposed study will use structured data fields only and will not involve chart review or validation to obtain additional information on the adverse events other than the study outcomes of interest.

14. Plans for Disseminating and Communicating Study Results

This study protocol will be registered in the ENCePP EU PAS Register[39] prior to the implementation of the diagnostic phase. In addition, study findings will be disseminated as manuscript(s) in peer-reviewed journals and/or as conference abstract presentations at international professional conferences.

15. References

1. NICE. NICE strategy 2021 to 2026 [Internet]. 2021 [cited 2021 Jun 24]. Available from: https://www.nice.org.uk/about/who-we-are/corporate-publications/the-nice-strategy-2021-to-2026
2. Wang SV, Pinheiro S, Hua W, Arlet P, Uyama Y, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. BMJ 2021;372:m4856
3. Ruff CT, Giugliano RP, Braunwald E, Hoffman EB, Deenadayalu N, et al. Comparison of the efficacy and safety of new oral anticoagulants with warfarin in patients with atrial fibrillation: a meta-analysis of randomised trials. Lancet. 2014;383(9921):955-62. Available from: https://pubmed.ncbi.nlm.nih.gov/24315724/. Accessed August 18, 2021
4. European Medicines Agency (EMA). Human Medicines European Public Assessment Report (EPAR). January 2021. Available from: https://www.ema.europa.eu/en/medicines. Accessed August 25, 2021.
5. Electronic Medicines Compendium (EMC). Authorized Medicines in the UK. June 2021. Available from: https://www.medicines.org.uk/emc/product/6907/smpc#AUTHDATE. Accessed August 25, 2021
6. National Institute for Health and Care Excellence (NICE). Atrial Fibrillation Management Guidelines. June 2014. Available from: https://www.nice.org.uk/guidance/cg180. Accessed August 25, 2021
7. National Institute for Health and Care Excellence (NICE). Atrial Fibrillation: Diagnosis and Management. April 2021. Available from: https://www.nice.org.uk/guidance/ng196. Accessed August 25, 2021
8. Veradigm. Using Veradigm's EHR Dataset to Replicate ROCKET-AF Clinical Trial [Internet]. 2021 [cited 2021 Jun 11]. Available from: https://veradigm.com/veradigm-news/ehr-dataset-to-replicate-rocket-af-clinical-trial/
9. Randomized Controlled Trials Duplicated Using Prospective Longitudinal Insurance Claims: Applying Techniques of Epidemiology (RCT DUPLICATE) [Internet]. RCT DUPLICATE. [cited 2020 Jul 15]. Available from: https://www.rctduplicate.org/
10. Electronic Medicines Compendium (EMC). Authorized Medicines in the UK. June 2021. Available from: https://www.medicines.org.uk/emc/product/6907/smpc#AUTHDATE. Accessed August 25, 2021
11. European Society of Cardiology (ESC). Atrial fibrillation set to affect more than 14 million over-65s in the EU by 2060. 6 June 2019. Available from: https://www.escardio.org/The-ESC/Press-Office/Press-releases/Atrial-fibrillation-set-to-affect-more-than-14-million-over-65s-in-the-EU-by-2060. Accessed 17 August 2021
12. National Health Service (NHS). Atrial Fibrillation. 17 May 2021. Available from: https://www.nhs.uk/conditions/atrial-fibrillation/. Accessed 17 August 2021
13. Burdett P, Lip GYH, Atrial fibrillation in the United Kingdom: predicting costs of an emerging epidemic recognising and forecasting the cost drivers of atrial fibrillation-related costs. Eur Heart J Qual Care Clin Outcomes. 2020;qcaa093. Available from: https://pubmed.ncbi.nlm.nih.gov/33346822/. Accessed August 18, 2021.
14. Fohtung RB, Rich MW, Identification of patients at risk of stroke from atrial fibrillation. US Cardiology Review. 2016;10(2):60-4. Available from: https://www.uscjournal.com/articles/identification-patients-risk-stroke-atrial-fibrillation. Accessed August 18, 2021.
15. Schneeweiss S, Gagne JJ, Patrick AR, Choudhry NK, Avorn J. Comparative efficacy and safety of new oral anticoagulants in patients with atrial fibrillation. Circ Cardiovasc Qual Outcomes. 2021;5(4):480-6. Available from: https://pubmed.ncbi.nlm.nih.gov/22787066/. Accessed August 18, 2021

16. Morimoto T, Crawford B, Wada K, Ueda S. Comparative efficacy and safety of novel oral anticoagulants in patients with atrial fibrillation: a network meta-analysis with the adjustment for the possible bias from open label studies. J Cardiol. 2015;66(6):466-74. Available from: https://pubmed.ncbi.nlm.nih.gov/26162944/. Accessed August 18, 2021
17. Hernan MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available: table 1. *Am. J. Epidemiol.* 183(8), 758–764 (2016).
18. Granger, CG, Alexander JH, McMurray JJV, Lopes RD, Hylek EM, et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2011; 365:981-992, DOI: 10.1056/NEJMoa1107039
19. Patel MR, Mahaffey KW, Garg J, Singer DE, Hacke W, et al. Rivaroxaban versus Warfarin in Nonvalvular Atrial Fibrillation. N Engl J Med 2011; 365:883-891. DOI: 10.1056/NEJMoa1009638
20. Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, et al. Exoxaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2013; 369:2093-2104. DOI: 10.1056/NEJMoa1310907
21. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, et al. Dibigatran versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2009; 361:1139-1151. DOI: 10.1056/NEJMoa0905561
22. Adderley NJ, Ryan R, Nirantharakumar K, Marshall T. Prevalence and treatment of atrial fibrillation in UK general practice from 2000 to 2016. Heart. 2019 Jan;105(1):27–33.
23. Ruigómez A, Vora P, Balabanova Y, Brobert G, Roberts L, Fatoba S, et al. Discontinuation of non-Vitamin K antagonist oral anticoagulants in patients with non-valvular atrial fibrillation: a population-based cohort study using primary care data from The Health Improvement Network in the UK. BMJ Open [Internet]. 2019 Oct 18 [cited 2021 Jun 11];9(10). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6803078/
24. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. PDS 2006, https://doi.org/10.1002/pds.1335.
25. Adderley NJ, Nirantharakumar K, Marshall T. Risk of stroke and transient ischaemic attack in patients with a diagnosis of resolved atrial fibrillation: retrospective cohort studies. BMJ 2018;361:k1717
26. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest. 2010 Feb;137(2):263–72.
27. Adderley NJ, Nirantharakumar K, Marshall T. Risk of stroke and transient ischaemic attack in patients with a diagnosis of resolved atrial fibrillation: retrospective cohort studies. BMJ 2018;361:k1717
28. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest. 2010 Feb;137(2):263–72.
29. Lane D, Lip GYH. Using the CHA2DS2-VASc and HAS-BLED scores to aid decision making for thromboprophylaxis in nonvalvular atrial fibrillation. Circulation. 2012;126:860–865.
30. O'Brian EC, Simon DN, Singer D, Thomas L, Hylek EM, et al. Abstract 18832: The ORBIT-AF bleed score: a simple score to assess major bleeding risk in atrial fibrillation. Circulation. 2014;130:A18832.

31. Hannon N, Sheehan O, Kelly L, Michael Marnane, et al. Stroke associated with atrial fibrillation -- Incidence and early outcomes in the North Dublin population stroke study. Cerebrovasc Dis. 2009;29(1):43-49.
32. Hernan MA. Causal analysis of existing database: no power calculations required. JClinEpi. In press Aug 28, 2021: https://doi.org/10.1016/j.jclinepi.2021.08.028
33. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009;28(25):3083-107
34. National Health Service (NHS). Quality and Outcomes Framework (QOF), enhanced services and core contract extraction specifications (business rules). Updated July 2021. Available from: https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/quality-and-outcomes-framework-qof. Accessed October 4, 2021.
35. Johnson ES, Bartman BA, Briesacher BA, et al. The incident user design in comparative effectiveness research. Pharmacoepidemiol Drug Saf. 2013;22(1):1-6.
36. Park Y, Bateman BT, Kim DH, et al. Use of haloperidol versus atypical antipsychotics and risk of in-hospital death in patients with acute myocardial infarction: cohort study. BMJ, 2018;360:k1218.
37. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Intention-to-treat versus per-protocol analysis. Perspect Clin Res. 2016;7(3):144-146.
38. Delaney JAC, Seeger JD. Sensitivity analysis, In: Developing a protocol for observational comparative effectiveness research: a user's guide. AHRQ Publication No. 12(13)-EHC009. ROchville, MD: Agency for Healthcare Research and Quality; January 2013: Chapter 11, pp. 145-60.
39. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP).The European Union electronic Register of Post-Authorisation Studies (EU PAS Register). Available from: http://www.encepp.eu/encepp/studiesDatabase.jsp. Accessed October 4, 2021.