Real-world effectiveness of oral anticoagulants in the prevention of stroke: emulation and extension of the ARISTOTLE trial using UK EHRs

Maud Emma Louise Teoh

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy of the University of London

November 2024

Department of Non-Communicable Disease Epidemiology

Faculty of Epidemiology and Population Health

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by Medical Research Council [grant number MR/N013638/1]

Research group affiliation: Electronic Health Records Research Group

Declaration

I, Maud Emma Louise Teoh, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: .....  .....

Date: 13 May 2024

Use of published work

One paper has been published and one paper is under review based on work undertaken for this thesis. Work for these papers were carried out as part of the PhD and took place during the period of registration for the PhD. For these papers, Maud Teoh was the lead author, and prepared all protocols and drafts of the papers. The contributions of the co-authors were restricted to providing study advice and comments on the drafts prepared by Maud Teoh (published under middle name/maiden name: Emma Maud Powell).

Abstract

Background

Stroke prevention treatment guidance for patients with atrial fibrillation (AF) uses evidence generated from randomised controlled trials (RCTs). However, applicability to patient groups excluded from the trials remains unknown. Real-world patient data provides an opportunity to evaluate outcomes in a trial analogous population of direct oral anticoagulants (DOACs) users and in patients otherwise excluded from RCTs, however there remains uncertainty on the validity of the methods and suitability of the data.

This thesis sought to validate non-interventional methodology for comparison of treatment effectiveness of oral anticoagulants in AF by emulating the pivotal ARISTOTLE trial (apixaban vs warfarin) in linked UK primary care data before extending the analysis to study groups excluded from, or underrepresented in ARISTOTLE.

Methods

This thesis used a novel method involving simultaneous equations and sampling to select a subset of patients with AF in CPRD Aurum prescribed apixaban or warfarin that matched the ARISTOTLE participants on baseline characteristics using only publicly available summaries. Recently developed methods for inclusion of prevalent users were explored, and a sampling method used with a modification to mimic the process of screening into an RCT. ARISTOTLE outcomes were assessed during 2.5 years of patient follow-up and results benchmarked before extending the analysis to patient groups under-represented in or excluded from ARISTOTLE.

Results

I was able to select a subset of patients in CPRD Aurum that matched ARISTOTLE participants on summary baseline characteristics and included prevalent users. The analysis sample comprised 8734 apixaban users and propensity-score matched 8734 warfarin users in CPRD. Results demonstrated non-inferiority of apixaban vs warfarin consistent with the pre-specified benchmarking criteria. Unlike in ARISTOTLE superiority of apixaban vs warfarin was not seen which may be linked to the higher proportion of patients with well-controlled warfarin and lower proportion of Asian patients compared to ARISTOTLE. After benchmarking results to ARISTOTLE, I extended the analysis to look at an under-represented group (people aged ≥ 75 years) and an excluded patient group (increased bleeding risk). In the people aged ≥ 75 years consistent results were seen compared with people aged < 75 years and with the ARISTOTLE emulation with similar risks of stroke/SE and all-cause death for apixaban vs warfarin along with a trend for a lower risk of major bleeding on apixaban compared with warfarin. The increased bleeding risk group also showed results consistent with the ARISTOTLE emulation for key outcomes.

Conclusions

Emulation of a reference trial in oral anticoagulants for atrial fibrillation can aid understanding of results in non-interventional data and increase confidence in the methods used facilitating the extension to patient groups of interest excluded or underrepresented in the trial. The framework can be adapted to investigate treatment effects for other conditions.

Acknowledgements

I am very grateful for the support and advice over many years from my primary supervisors, Kevin Wing and Ian Douglas, for all the good ideas they shared, their wealth of experience and knowledge and for guiding me on research methods. I am especially thankful for the encouragement and friendliness in every meeting. I would like to thank my additional supervisor Usha Gungabissoon for her advice on the project, for sharing her extensive experience in this field, and for always making me feel tasks were achievable.

I would like to thank John Tazare for his guidance on prevalent new user design and alternative methods that helped guide methods used in this work. I would like to thank Angel Wong and Paris Baptiste for reviewing my results paper and providing helpful suggestions. I would like to thank Liam Smeeth for his review and the EHR group at LSHTM for being such a welcoming group. I would like to thank Turki Bin Hammad and Lin Yan Ling for making use of my datasets to explore other questions.

I would like to thank the Medical Research Council for funding this PhD and funding for training courses to develop my skills. I would like to thank the MRC LID team for their support, and the training and events they organized.

I would like to thank my husband Jeremy for his support over the years and my daughters Ami and Amber for their patience and for bringing me so much happiness. I would like to thank my parents Andrew and Florence for their support.

Contents

List of Tables

List of Figures

List of Abbreviations

ACEi	Angiotensin converting enzyme inhibitor
AF	Atrial fibrillation
ARB	Angiotensin II receptor blocker
BMI	Body mass index
CEM	Coarsened exact matching
CI	Confidence interval
CKD	Chronic kidney disease
CPRD	Clinical Practice Research Datalink
DOAC	Directly acting oral anticoagulants
DVT	Deep vein thrombosis
ECG	Electrocardiogram
EHR	Electronic health records
EMA	European Medicines Agency
FDA	Food and Drug Administration
GP	General practitioner
HES	Hospital Episode Statistics
HR	Hazard ratio
ICD-10	International Classification of Diseases, 10th Revision
ICH	Intracranial haemorrhage
IMD	Index of multiple deprivation
INR	International normalised ratio
IPTW	Inverse probability of treatment weighting
IQR	Interquartile range
LVEF	Left ventricular ejection fraction
MHRA	Medicines and Healthcare products Regulatory Agency
MI	Myocardial infarction
NICE	National Institute for Health and Care Excellence
NVAF	Non-valvular atrial fibrillation
OAC	Oral anticoagulant
ONS	Office for National Statistics
PE	Pulmonary embolism
PNU	Prevalent new user
PSM	Propensity score matching
PT	Prothrombin time
RCT	Randomised controlled trial
RWE	Real world evidence
Rx	Prescription
SE	Systemic embolism
TIA	Transient ischemic attack
TTR	Time in therapeutic range
UK	United Kingdom
VKA	Vitamin k antagonist
VTE	Venous thromboembolism

Chapter 1 Introduction

1.1. Rationale

Drugs are licensed based on a favourable benefit risk profile from randomised controlled trials (RCT)s in terms of its efficacy and safety in a controlled setting. Whilst RCTs are necessary in proving the effects of the drug, they have several limitations. RCTs typically include a highly selected patient population to minimise the risk of harm (i.e. unintended drug effects) and thus will often include the healthiest subset of the target patient population for example by including only younger patients, incident disease, or those without significant comorbidities or concomitant therapies by applying strict eligibility criteria to exclude individuals based on comorbidity profile. In some instances, RCTs will include an enriched population to maximise the chance of observing a statistically significant treatment effect, thus trial populations may not be representative of the indicated (real-world) population. Furthermore, trials are typically of short duration and may therefore not include patients exposed long-term to the medication or detect long-term outcomes; the limited sample size and follow-up of RCTs also means rare adverse events may not be observed. As a result, evidence on treatment benefits and harms is often lacking for patients who would not have met the eligibility criteria of the RCT, or who are prescribed drugs in a different way than administered in a trial setting:

- A study of patients with inflammatory bowel diseases found only 31% would have met the inclusion criteria of RCTs in biological agents. Reasons for exclusion included taking high doses of steroids and comorbidities (2).
- A study of 'real life' Multiple Sclerosis patients found mean time since first symptoms to first-line disease modifying therapy treatment initiation was approximately 4 years shorter than for RCT participants (3).

This poses a problem for the clinician post-licensure in that they must extrapolate RCT results from a highly selected population to their patient case which may not be appropriate. The patient may differ from the trial participants in ways such as: having comorbid conditions that were in the trial exclusion criteria, being on other medications prohibited by the RCT, being in an age group under-represented or excluded from the trial, or by a difference in the severity or treatment history for the indication of interest. There have been cases where drug effectiveness has appeared to differ in certain patient subpopulations; even if relative risks do not differ across subgroups absolute risks often do which can have an impact on risk benefit assessment, for example in a subgroup with lower absolute risks of an effectiveness outcome the benefit in the reduction in risk of that outcome may be outweighed by the harms associated with a given medication. For example, pivotal RCTs for antihypertensives included patient populations that mostly had more severe hypertension for secondary prevention. Post-licensure, antihypertensives were being prescribed to patients with mild hypertension for primary prevention despite this patient group being excluded or under-represented in the RCTs. The Cochrane meta-analysis of RCT data for this patient group had a relatively small sample size (N=8912) and found antihypertensive drugs were not effective in the treatment of adults with mild hypertension for primary prevention (4).

In the post authorisation setting in the US, The 21st Century Cures Act passed in 2016 places an increased focus on the use of data collected as part of routine care to support regulatory decision making. The European Medicines Agency (EMA) can also require post authorisation safety and effectiveness studies in wider populations (5, 6). Non-interventional (sometimes also referred to as “observational” or “real world”) data sources overcome many of the RCT limitations given that they contain data for a diverse range of patients treated with the drug in routine care including patients who would have been not eligible for the trials. Furthermore, non-interventional data allows the study of adherence to treatment in routine care which is of

interest given patients may be less adherent outside the setting of a clinical trial. Drug effectiveness and safety out of the controlled trial setting where patients may have more comorbidities and concomitant medications can also be explored, along with the study of longer-term clinical outcomes with the drugs being used for longer durations than in the RCTs or used for a similar duration but with longer follow-up available in the non-interventional data.

Data collected as a part of routine patient care such as diagnoses and prescription data recorded in electronic healthcare record (EHRs) provide a valuable opportunity to obtain evidence on drug effectiveness in a routine care setting. These data can provide far larger sample sizes compared with the numbers that are included in RCTs and from a more diverse population. A wide breadth of outcomes including rare outcomes may be captured in EHRs along with detailed medical history and sociodemographic and lifestyle data.

A key problem with non-interventional studies using these data is that the absence of randomisation leaves them highly susceptible to confounding, making it difficult to have confidence in the results. Confounding by indication is a particular problem in observational research(7). Particular care must also be taken in attempting to determine drug effectiveness (as compared with analysis of comparative safety where confounding by indication is generally less of a problem) because of the risk of bias, including several forms of selection bias either inherent in the data (such as channelling bias in the clinician selection of treatments for patients or attrition bias in terms of systematic differences in patients switching or stopping treatments during follow-up) or bias introduced in suboptimal design of the study for example in the selection of index date for prevalent users.

One approach to try and reduce some of the inherent uncertainties with the analysis of non-interventional data is to match patient records in EHRs to those from an existing RCT (or

'reference trial') followed by matching within EHR treatment groups to select an EHR population similar to the trial population that is well balanced by treatment group. If the estimates of effectiveness and safety obtained from this approach are comparable with the trial results then this provides confidence in the validity of the non-interventional results obtained and by extensions the methods used.

If non-interventional data can be successfully used to approximate the findings of an RCT in this way then the analysis can be extended to estimate the effects in groups underrepresented, or excluded from the original (reference) RCT. The ability to emulate a reference trial and the optimal methods to use to do so are likely to vary by therapy area due to differing confounders and different treatment patterns during follow-up. To emulate a reference trial using non-interventional data, a suitable trial amenable to emulation in EHR data must first be identified; namely a treatment that is recorded (eg as prescriptions in EHRs or insurance claims) and with valid records of diagnoses and outcomes. This project will initially involve testing whether an RCT can be successfully emulated in EHR data while developing optimal methodology for that drug/therapeutic area. This methodology can then be applied to generate evidence on groups that were underrepresented in the original (reference) RCT.

1.2. Aim

To investigate the use of United Kingdom (UK) EHRs in determining effectiveness of oral anticoagulants for the prevention of stroke in atrial fibrillation (AF) through emulating a reference trial and benchmarking the results in the trial-analogous EHR cohort against the reference trial results. Subsequently analysis will be extended to groups underrepresented in or excluded from the reference trial.

1.3. Objectives

The aim will be addressed by the following objectives, specifically to:

1. Emulate the reference trial ARISTOTLE comparing apixaban to warfarin for prevention of stroke in atrial fibrillation in UK EHRs including application of the trial eligibility criteria, matching to the baseline characteristics of the participants in the reference trial, and assessing the validity of the results and methods by benchmarking.
2. To explore different methods in the emulation of the reference trial including different methods of matching and the inclusion of prevalent users.
3. To use the methodological framework to extend the analysis to look at apixaban compared to warfarin in patient groups with atrial fibrillation underrepresented in or excluded from the reference trial

1.4. Thesis structure

This thesis is a mixed book and research paper style containing both chapters based on journal articles along with traditional thesis style chapters.

Chapter 1 introduces the rationale for the PhD along with the aim and objectives.

Chapter 2 describes the background and feasibility work involved in selection of a target RCT including: i) a brief summary of recent approaches to emulation (and extension) of reference trials in non-interventional research ii) the data sources and how these relate to the aim of the PhD, iii) a summary of feasibility work and the selection of the target RCT ARISTOTLE, and iv) background information on atrial fibrillation, treatment for atrial fibrillation, and key questions on the applicability of ARISTOTLE results to the UK population with this indication.

Chapter 3 summarises the results of a literature review conducted to assess use of EHRs in assessing drug effectiveness in prevention of stroke in patients with AF with a focus on the

methods used and degree to which the studies attempted to emulate and extend an existing RCT.

Chapter 4 presents a protocol paper outlining the planned methods and analyses for the trial emulation study published in the *BMJ Open* (1) including how the trial inclusion and exclusion criteria were applied, and determination of the benchmarking criteria. The protocol paper is followed by a more detailed description of the development and selection of methods used in constructing a trial-analogous cohort, including selection of EHR patients matching the trial participants, and the method for inclusion of prevalent users.

Chapter 5 is a results paper covering the main effectiveness and safety results in the emulation of ARISTOTLE in Clinical Practice Research Datalink (CPRD) Aurum study under final review in May 2024 for publication in *PLOS Medicine* and additional detail on the results including the benchmarking and comparison against the RCT results , sensitivity analyses, and an analysis of results in the new users of apixaban and warfarin that were eligible for ARISTOTLE (without matching to the trial participants).

Chapter 6 summarises the results in special patient populations that were under-represented in ARISTOTLE (elderly patients) or excluded from ARISTOTLE (patients with increased bleeding risk).

Chapter 7 concludes the thesis with a discussion of the findings of the analyses, the strengths and limitations of the work and ideas for possible future work.

Chapter 2 Background

This chapter will present:

- A brief background to the topic of trial emulation including a description of past studies and commonly used methods.
- The decision framework and feasibility assessment performed in selecting a reference trial that could be emulated using UK EHRs
- A background to the therapeutic area of interest for the selected reference trial, atrial fibrillation.
- A description of the treatments available for stroke prevention in atrial fibrillation and the pivotal RCTs for these treatments in this indication.
- A summary of the selected reference trial for emulation, ARISTOTLE, which compared apixaban to warfarin for prevention of stroke or systemic embolism in patients with atrial fibrillation.

Chapter 1 established the rationale of this PhD, namely the increasing availability of routinely collected medical data creating an opportunity for looking at effectiveness of treatments in a 'real world' setting and the potential for 'emulation' of a reference trial to validate methods used in estimating treatment effectiveness using non-interventional data. These non-interventional data provide the potential to assess the safety and effectiveness of treatments without the usual extra administration required in post-marketing or Phase 4 studies; furthermore the patients prescribed treatments of interest in the routine care setting are likely to include patients that would not have been eligible to be included in the RCTs for the treatment or to be from relatively underrepresented demographics. This led to the additional rationale of the PhD that following the successful emulation of a reference trial, the validated methods could be used to look at treatment effectiveness in excluded or underrepresented

patient groups. Analysis of safety was not the focus of this thesis though safety outcomes analysed by the reference trial can also be assessed during the emulation.

This PhD did not start with a specific reference trial in mind and was instead open to the possibility of emulating any reference trial using UK EHRs. This meant the first step was selection of a suitable reference trial to emulate. In order to assess the feasibility of emulation of a potential reference trial key restrictions come from the limitations of the data available in the planned data source of UK EHRs.

2.1. Trial Emulation

Whereas target trial emulation is attempting to emulate a hypothetical RCT that does not exist, reference trial emulation (also called RCT replication or benchmarking, the focus of this thesis), is trying to emulate an existing RCT.

2.1.1. Target trial emulation

Target trial emulation describes the application of the design and methods used in RCTs to observational data with the aim of measuring the effect of an intervention imagining a hypothetical RCT as the target trial, for example looking at the risk of dementia with the use of proton pump inhibitors (8) or suicide-related events in antipsychotics in post-traumatic stress disorder patients(9). Hernan et al introduce the idea of emulating a target trial in 2016 (10) noting the numerous advantages of the approach: facilitating the use of causal inference methods that use counterfactual reasoning, making a link between methods used for observational studies and RCTs, preventing common biases in observational studies, and providing estimates that can be more easily compared between different target trial emulations in contrast to comparing estimates from traditional non-interventional studies..

The use of the target trial emulation approach helps researchers avoid different sources of bias, increases the rigour of the study by adopted principles common in RCTs such as finalising a protocol or statistical analysis plan prior to analyses, and helps standardise the reporting of the results. Hernan et al (11) described how this framework can avoid immortal time bias and other biases.

2.1.2. Confounding

A key challenge in emulating a reference trial, and for the area of trial emulation more generally, is how to remove the effect of confounding from treatment effect estimates. Confounding is defined as bias in the estimation of the effects of an exposure on an outcome due to inherent differences in risk between exposed and unexposed individuals (12). In an RCT, the process of randomisation removes the link between a patient's baseline condition or prognosis and the choice of treatment meaning the treatment effect estimates obtained are unbiased estimates of the average treatment effect (13) (assuming an appropriate estimand strategy has been employed). Routinely collected healthcare data lacks this randomisation meaning there is a high likelihood the probability of a treatment being given to a patient depends on factors such as their baseline characteristics, disease history, or response to past treatments. This problem can be particularly acute in the situation of newly available treatments, so-called 'channelling bias' (14), where there may be a systematic channelling of a particular subset of the target population (for example a healthier subset of the patients, patients that have failed prior treatments, or patients with a more severe disease state) with the nature of the subset depending on the guidance issued to the prescribers and/or rules from the payers.

There are a range of methods that can be used in an attempt to deal with the confounding that results from the lack of randomisation including forms of 'matching', inverse probability of

treatment weighting, and adjusting for the baseline variables in the statistical model of the outcomes to produce adjusted treatment estimates (15). The baseline variables that should be considered in these methods depends on the treatment and therapeutic area; prior expert knowledge can help inform an appropriate choice and the variables displayed in the 'baseline characteristics table' of a published RCT serve as a logical initial list including variables such as age, sex, disease severity measures, relevant comorbidities, and concomitant medications.

Matching methods involve matching pairs of patients in each treatment arm of interest on their baseline variables. This can be done via 'exact matching' (16) in which subgroups of patients with identical combinations of covariates are matched (for example identical sex, disease severity, and age group), or via propensity score matching. The concept of the propensity score was introduced by Rosenbaum and Rubin(17) and is defined as the probability of treatment assignment conditional on observed baseline covariates. The propensity score can be estimated by regressing treatment on the measured baseline characteristics most commonly by using a logistic regression model. The propensity score functions as a balancing variable since conditional on the propensity score, the distribution of measured baseline covariates will be similar between the patients in the different treatment arms.

The propensity score greatly simplifies the task of matching patients in that patients can be matched on a single number rather than attempting to simultaneously match on multiple covariates with different matching algorithms available (18). Through the process of propensity score matching (PSM) a subset of patients in the cohort of interest is selected which should be well balanced on all baseline covariates included in the propensity score model.

Inverse probability of treatment weighting (IPTW) introduced by Rosenbaum(19) is also commonly used in trial emulation and involves constructing weights based on the inverse of a patient's propensity score. Imagining a hypothetical observational study comparing a new treatment to a comparator for treatment-resistant depression; if the comparator group had more patients with a high number of prior treatment failures (predictive of future treatment success) then these patients could be down-weighted to create a pseudo-population matching the population of patients on the new treatment thereby allowing valid treatment effects to be estimated. The method differs from propensity score matching in that instead of dropping patients that are unmatched from the cohort, all patients are instead kept in the cohort. By applying the weights a 'pseudo-population' is produced which should be well-matched on the baseline covariates included in the propensity score model.

A key appeal of PSM and IPTW is the avoidance of having to include the baseline covariates in the modelling of the outcome measure. An alternative method to adjust for confounding is by including the baseline covariates directly in the statistical model for the outcome of interest, producing 'adjusted' treatment estimates. As the number of baseline covariates to be included increases the modelling can become complex and treatment estimates less easy to interpret.

2.1.3. Reference trial emulation

A special case of target trial emulation involves the benchmarking against a real trial, termed 'reference trial emulation' or 'RCT replication'. Benchmarking could also be achieved without using the target trial emulation framework. Attempting to emulate a reference trial (a real completed RCT) with publicly available results has the benefit of allowing the researcher to pre-specify benchmarking (validation) criteria based on the reference trial results. A potential disadvantage of this approach is that this may bias the researcher into modifying or

manipulating their analysis to obtain results close to the reference trial results or meeting the benchmarking criteria. Other researchers (RCT-DUPLICATE) have also attempted to emulate ongoing reference trials (20) taking advantage of their being blind to results that do not yet exist to avoid this potential source of bias. The risk for researcher bias can be mitigated by publishing a protocol or statistical analysis plan in advance of the analysis.

The introduction chapter of the thesis briefly introduced some reasons why the emulation of a reference trial may be of interest, most importantly to serve as validation of the data sources and methods in the therapeutic area and help improve and guide future trial emulation work in the area. Future emulation work may cover emulation of other historical or future reference trials or emulation of target trials for patient groups or research questions that cannot be studied in an RCT whether that be for ethical or logistical reasons.

Similar to the early days of RCTs, there are still many unknowns in the best approaches to use when looking at questions of treatment effectiveness in observational data. As the body of evidence in this area increases in size we might expect to see methodological advances in the future with regulatory and funding bodies eventually able to provide more detailed guidance on recommended methods and design aspects analogous to the level of detail in guidance available today for RCTs.

The Food and Drug Administration (FDA) framework on real world evidence (RWE) published in 2018 notes reference trial emulation '*may provide insight into the opportunities and limitations of using these designs in regulatory decisions'.* Whilst this thesis will focus on the emulation of one reference trial using UK EHR data, this will play a part in adding to the body of evidence on the data sources and methods that can help guide improved design of future trial emulation studies and of how this data can be used to answer questions not answered by the RCTs.

The emulation of a reference trial using UK EHRs may help elucidate the question of generalisability of results from a multi-centred RCT to the UK population.

2.1.4. Past reference trial emulation studies

A literature review on the emulation of reference trials using real world data was performed by Baptiste in Feb 2023 (21). As part of the review, nine studies (22-31) involving emulation of a reference trial were identified. This included one study, from RCT-DUPLICATE (22), which represents a large-scale initiative to emulate multiple reference trials. Additional studies published since the 2023 review of particular relevance include the emulation of an antihypertensive trial(32) and further results from the RCT-DUPLICATE initiative (20). The identified studies emulated reference trials in a range of therapeutic areas, including: psoriasis, cancer, diabetes, myocardial infarction (MI), chronic obstructive pulmonary disease (COPD), suicidality, hypertension, atrial fibrillation, osteoporosis, chronic kidney disease (CKD), heart failure, and asthma treatments.

Most of the identified studies emulated the reference trial using US claims data (including all RCT-DUPLICATE studies) other than 2 studies that used Swedish registry data, 1 study that used linked administrative databases from Canada, 1 study that used a UK pharmacovigilance register of patients, and 2 studies that used UK electronic healthcare records - Wing et al (33) looking at treatments for COPD emulating the TORCH RCT, and Baptiste et al (32) looking at antihypertensives emulating the ONTARGET RCT.

2.1.5. Reference trial emulation methods

The key methodological choice of how to deal with confounding has been described earlier in this chapter (Section 2.1.2) with most researchers using methods such as exact matching, PSM, IPTW, or multivariable adjustment of treatment estimates. Existing reference trial emulation studies employed a mix of the described methods with propensity score matching being the most commonly used (all of the RCT-DUPLICATE studies) followed by inverse probability of treatment weighting.

Other study design choices can impact the results obtained though there are approaches common across all studies, namely the application of the RCT eligibility criteria (the trial inclusion and exclusion criteria) to the cohort of real world evidence (RWE) patients. By applying the eligibility criteria, the researcher aims to select the subset of the RWE population that would have been eligible to participate in the reference trial. Applying the RCT eligibility criteria alone may not result in a RWE cohort that matches the reference trial population on baseline covariates; the RCT-DUPLICATE authors noted that important patient characteristics including age, sex, and comorbidities often differed between their RWE cohorts and the reference trials (20).

Two studies, the emulation of TORCH by Wing et al (33) and the emulation of ONTARGET by Baptiste et al (21), included an additional step in the emulation process in which after applying the trial eligibility criteria they attempted to match the RCT participants on baseline characteristics. Both studies matched to the RCT participants by propensity score matching using individual patient data from the target reference trial. This additional step was effective at selecting an 'RCT-analogous' cohort of patients that not only met the trial eligibility criteria but looked similar to the trial population on key baseline characteristics. Wing et al results showed that prior to the matching step, the RWE cohort differed to the TORCH

participants on sex (62% male vs 76% male in TORCH), history of cardiovascular disease (28% vs 51%), and lung function measured as FEV_1 % of predicted (51.7 vs 44.2); after the matching step all these baseline characteristics moved closer to the trial participants (76% male, 46% history of cardiovascular disease, 47.2 lung function) (33).

An additional consideration in the emulation of reference trials is whether to include prevalent users of the treatments of interest vs restricting the RWE cohort to include new users only. The design of the reference trial may include prevalent users of a treatment, either exclusively or a certain proportion of participants. Inclusion of prevalent users in observational studies is challenging given the risk of introducing selection bias; the majority of existing reference trial emulation studies to date have restricted the RWE cohort to new users (for example all 32 of RCT-DUPLICATE studies were in new users (20)) whilst the Baptiste replication of ONTARGET and the Wing emulation of TORCH did include prevalent users (21, 25).

2.1.6. Benchmarking/validation of reference trial emulation studies

A key question in the field of reference trial emulation surrounds the comparison of the results obtained in the real-world cohort with the reference trial results. What criteria should be used to determine if the results are equivalent and how should any difference obtained be assessed? In her review Baptiste noted that all studies identified in her search benchmarked their results against the reference trial results though not all studies pre-specified their benchmarking criteria (21); the more recently identified studies of the ONTARGET emulation in CPRD and the RCT-DUPLICATE studies all prespecified benchmarking criteria.

In a comparison of 32 RCT emulations(34) RCT-DUPLICATE planned 3 binary agreement measures : (i) ‘full statistical significance agreement’ defined as the RWE study replicating

both the direction and statistical significance of the RCT results; (ii) 'estimate agreement' where a RWE hazard ratio (HR) estimate was within the 95% CI of the RCT estimate; (iii) hypothesis tests involving calculating the standardised difference between the RCT and RWE estimates to determine whether the RWE treatment effect estimates were different to the RCT findings (with $P < 0.05$ considered statistically significant). The RCT-DUPLICATE authors also defined a weaker criterion 'partial significance agreement' for cases where the RWE study met the RCT noninferiority criteria but not superiority for replication of non-inferiority reference trials that demonstrated superiority.

Baptiste notes that all studies identified were able to replicate some of the results (21), the more recent replication of ONTARGET in CPRD was also able to replicate the RCT results (21). A study by RCT-DUPLICATE, assessing RCT-RWE concordance of 32 RCTs they had emulated, found 75% of the studies met the criteria for statistical significance agreement, 66% estimate agreement, and 72% standardised difference agreement (20). A post-hoc analysis performed by RCT-DUPLICATE looking at only the subset of 16 studies that had closer emulation of the reference trial found higher rates of successfully matching of the RCT results (with 94% meeting statistical significance, 88% estimate agreement, and 88% standardized difference agreement) (20). The RCT-DUPLICATE authors concluded that "Emulation differences, chance, and residual confounding can contribute to divergence in results and are difficult to disentangle"(20).

A recent publication from RCT-DUPLICATE by Heyard et al (34) performed a meta-analysis of 32 RCT emulations performed by the group to explore the sources of differences between RCT and RWE results. In this study 29 RCTs in which the primary outcome was a hazard ratio were selected with the data sources involving US insurance claims and Medicare data. A plot of the RWE HRs against the RCT HR was provided with the authors describing how the

approximately even scattering around the diagonal line showed an absence of systematic bias unlike other studies in which smaller effect sizes than the RCT tend to be seen in the RWE termed ‘shrinkage of effect size’(35).

2.1.7. Extensions of previous reference trial emulation studies

Only a few of the reference trial emulation studies extended their analyses to look at underrepresented or excluded patients groups. The TORCH emulation in patients with COPD in UK EHR was followed by an extension (25) in which patients that would been ineligible for the TORCH trial because of age, asthma, comorbidity or mild disease were included in the analysis. For the outcome of COPD exacerbations Wing found results were broadly consistent in the excluded patient groups with the exception of those with mild disease, in which a stronger protective association for FP-SAL compared with salmeterol was observed (risk ratio 0.56, 95% CI 0.46 to 0.70, vs. TORCH trial risk ratio 0.85, 95% CI 0.74 to 0.97). This study also detected an increased risk of mortality for FP-SAL vs salmeterol in those with prior asthma (hazard ratio 1.49, 95% CI 1.21 to 1.85, vs. TORCH trial-analogous HR 0.93, 95% CI 0.64 to 1.32).

Matthews et al emulated the VALIDATE trial comparing bilvalirudin to heparin during percutaneous coronary intervention for the outcomes of death, myocardial infarction, and bleeding using Swedish data, before applying the same framework to a target trial using data from the time period before the VALIDATE trial took place (26). Matthews found similar results in the target trial and reference trial emulation. In a separate study Matthews successfully emulated the TASTE trial (thrombus aspiring in ST-elevation myocardial infarction) using Swedish data before extending the analysis to assess the impact of sex and age group which found results consistent with the main analyses (27).

The successful emulation of ONTARGET by Baptiste in UK EHR was followed by exploring treatment effect heterogeneity of angiotensin converting enzyme inhibitor (ACEi) and angiotensin II receptor blocker (ARB) in groups under-represented in the reference trial: female patients, those aged ≥ 75 years and those with CKD (32) which observed similar treatment effects in all groups. Outcomes in different ethnic groups in England were also assessed using the same methodological framework (36) with Baptiste detecting no evidence of treatment effect heterogeneity for the primary outcome, whereas ARBs vs ACEi were associated with an increased risk of death for cardiovascular-related death in Black patients in contrast to white patients that saw a lower risk on ARBs vs ACEi. This study also detected higher absolute risks of angioedema in Black patients.

2.1.8. Previous reference trial emulation studies conclusion

Overall, the majority of the previous reference trial emulation studies were successful with most finding results consistent with their benchmarking criteria enabling in some cases the extension of the analysis to answer questions not answered by the reference trial with more confidence. The largest initiative, RCT-DUPLICATE, has published studies to date that used US insurance claims data meaning the findings of the initiative may not apply to other countries and data sources; in this thesis the emulation of a reference trial using UK electronic healthcare records should help explore the differences in reference trial emulation between US insurance claims and UK EHR. Most of the past reference trial emulation studies did not match to the baseline characteristics of the RCT participants or include prevalent users making these 2 additional areas worth exploring in this thesis.

2.2. Selection of a reference trial

2.2.1. Setting

The primary source of data for this project were UK electronic healthcare records (EHRs). In the UK, patients attend a General practitioner (GP) for most primary healthcare needs with

key data such as demographics, symptoms, laboratory results, vital signs, diagnosis, prescriptions, and attendance at accident and emergency hospital departments recorded electronically. Several different datasets of UK EHR exist with the sources that were planned to be used in this project described below.

Data source 1: CPRD Aurum

Clinical Practice Research Database (CPRD) Aurum is a database containing anonymised data from over 19 million patients with 7.1 million active patients as of September 2018 (13% of the population of England)(37). The data come from GP practices in England using EMIS software with diagnoses entered using a standardised international coding system (SNOMED CT). CPRD Aurum contains diagnoses, symptoms, lifestyle factors, prescriptions, referrals and tests and has been linked to national secondary care databases, deprivation data, and death registration data. CPRD Aurum is representative of the English population in age distribution, sex, geographical spread and social deprivation(37).

Data source 2: CPRD Gold

CPRD Gold contains primary care records similar to CPRD Aurum but based on practices using Vision software. CPRD Gold contains EHRs for over 5 million active patients. CPRD Gold is representative of the UK population with respect to age, gender and ethnicity(38). Diagnoses are entered into CPRD using READ codes, a hierarchical coding system[1].

Data source 3: THIN

The Health Improvement Network (THIN) contains medical records of 11.1 million patients (3.7 million active patients) collected from 562 general practices in the UK, covering 6.2% of the UK population. Overlap of patients in THIN and CPRD Gold is approximately 60% (39).

Data source 4: HES and ONS

[1] The UK is transitioning to using a standardised international coding system (SNOMED CT) in place of READ.

Information from the Hospital Episodes Statistics (HES) database was planned to be used to improve detection of the outcome events. A subset of CPRD Aurum and Gold contributing practices in England have patient records linked to HES, an administrative data source that contains patient demographic and diagnostic information, coded using the World Health Organisation International Classification of Diseases, 10th Revision (ICD-10) coding system, for every NHS hospital admission in England. All CPRD Aurum practices have been linked to HES whereas only a subset of CPRD Gold (56%) contributing practices have been linked to HES. CPRD records can also be linked to Office of National Statistics (ONS) mortality records, which provide information on cause of death for patients who die following a hospital referral (also coded using ICD-10).

Data source 5: RCT individual patient data requested via clinicalstudydatarequest.com and individual pharmaceutical/biotechnology company request portals

The website clinicalstudydatarequest.com(40) provides a database that can be used to identify RCTs that have made individual patient data available to researchers, typically following submission and approval of a research protocol. The list of RCTs on the clinicalstudydatarequest.com website was used as a starting list to guide the choice of reference trial for emulation followed by also checking the availability of patient data from pharmaceutical company websites.

The UK EHR data sources provide a wealth of patient data including detailed demographics, diagnoses, symptoms, regular lab results and vital signs, GP prescriptions, and certain outcomes and medical history from hospital data. The key data that were lacking from the UK EHR data at the time of selecting the reference trial to emulate were the prescriptions issued in secondary care, and other data relating to secondary care that may not be integrated into the patient's primary care EHR such as scan results, specialist test results, extra

information on diagnoses such as disease subtype, and genotyping of diseases. In considering feasibility it was therefore important to consider for each disease or condition and associated treatments of interest, whether it was a condition treated by GPs and whether the treatment can be prescribed by GPs.

Starting with this restriction (the reference trial must involve treatments prescribed by GPs) other criteria were added by considering the necessary data for emulation of a reference trial to be feasible.

2.2.2. Criteria for selection of a reference trial

The first key step of the project was selection of a suitable RCT which could be emulated within UK EHR data. For the RCT to be feasible it needed to have the following properties:

Table 2.1 RCT selection criteria

Criteria	Condition
1	RCT with published protocol and results, relevant to current UK practice (ie involving treatments currently recommended by NICE)
2	Prescription data available in UK EHRs (ie RCT includes only treatment(s) commonly prescribed by GPs)
3	Sample size – sufficient number of patients prescribed the active drug and any comparator(s) used in the RCT (number of patients prescribed the drug > number of patients randomised to the drug in the RCT)
4	Disease diagnosis (indication for the treatment) well recorded in EHRs
5	RCT outcomes well recorded in EHRs and/or linked HES and ONS mortality data
6	Subpopulations of interest prescribed the drug present in routine care who were excluded from or underrepresented in the RCT

EHR=electronic healthcare record; GP=General practitioner; HES=Hospital Episode Statistics; NICE= National Institute for Health and Care Excellence; ONS=Office for National Statistics; RCT=randomised controlled trial; UK=United Kingdom.

The Clinical Study Data Request database of RCTs (40) which had been marked as having data available for access by researchers was used as a starting point for consideration and the RCTs were systematically considered after grouping by therapeutic area and further by disease and drug. At this stage no decision had been made about a suitable

disease/therapeutic area; the following list of therapeutic areas were considered (examples of indications in parentheses):

- Neurology (epilepsy, multiple sclerosis)
- Psychiatry (depression, schizophrenia, anxiety)
- Oncology
- Diabetes
- Cardiovascular (statins, antihypertensives, oral anticoagulants, antiplatelets)
- Infectious diseases (antibiotics, vaccines, antivirals)
- Other (autoimmune conditions such as inflammatory bowel conditions, asthma, arthritis)

Many RCTs were ruled out on account of the treatments under study not being prescribed by a GP, such as chemotherapy and other treatments for cancer and biologic treatment for autoimmune and neurology conditions.

Some disease areas failed on criteria 4 requiring accurate diagnosis be captured in the EHRs, for example whilst epilepsy has treatments commonly prescribed by GPs after initiation by a neurologist, the type of epilepsy (both underlying cause and type of seizure) is a key selection criterion for epilepsy RCTs and was found to be not well recorded in EHRs.

Some common conditions such as depression have treatments prescribed by GPs and have a high number of patients exposed in EHRs (ie meeting conditions 1 through 4) but were not likely to have trial outcomes well recorded in EHRs given that the most commonly used primary outcomes in RCTs involved subjective depression symptom scales recorded at protocol-specified timepoints. In the UK EHRs it was judged unlikely for measures such as depression severity to be recorded at consistent timepoints post-treatment initiation for different patients due to the nature of non-interventional data collection; this would lead to

problems with missing data in any EHR cohort. Furthermore, the reason for a GP performing a questionnaire or recording a symptom would possibly be related to the disease severity or a change in symptoms causing problems in modelling data and making comparisons across patients.

More generally, RCTs involving longitudinal data typically employ different estimand ('what is to be estimated') strategies to estimate the treatment effect under different assumptions focused on the modelling techniques and/or imputation of data observed after an intercurrent event. Whilst certain intercurrent events would be expected to be well-recorded in EHRs (for example switching treatment and hospitalisation should be captured) the volume of missing data when compared with RCT scheduled timepoints and probable lack of independence between data being recorded and the underlying patient condition meant the risk of obtaining unreliable or biased treatment estimates in such an RCT replication was judged to be too high. In contrast 'hard' endpoints captured as a distinct clinical event are easier to emulate with EHRs since there is a much lower risk of missing data and in the case of time-to-event endpoints, the RCT estimand strategy is more likely to have a natural translation to differing censoring strategies in the EHR setting.

After applying the 6 conditions listed above the strongest candidates were RCTs involving relatively common conditions such as diabetes, hypertension, and atrial fibrillation which are commonly managed by GPs, have diagnosis well recorded, have 'hard' outcomes (such as stroke, cardiovascular events, and death) well recorded, and have subpopulations of interest prescribed the treatments in routine care.

At the time of the feasibility assessment, many RCTs into directly-acting oral anticoagulants (DOACs) for prevention of stroke in patients with atrial fibrillation and additional stroke risk factors had recently been published and had trial data available. Apixaban was selected for

investigation as a suitable RCT to replicate given it was a treatment recommended by National Institute for Health and Care Excellence (NICE) (41) and with a trend over time of increasing prescriptions (Figure 2.1).

The pivotal RCT of apixaban for stroke prevention in AF was 'The Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation' (ARISTOTLE)(42), which was designed for non-inferiority and demonstrated superiority of apixaban compared with warfarin in preventing stroke and systemic embolism (SE) in patients with nonvalvular AF and at least one risk factor for stroke. In ARISTOTLE apixaban was also shown to be superior to warfarin in all-cause mortality and the key safety outcome of major bleeding. Further detail on ARISTOTLE is given in section 2.5.

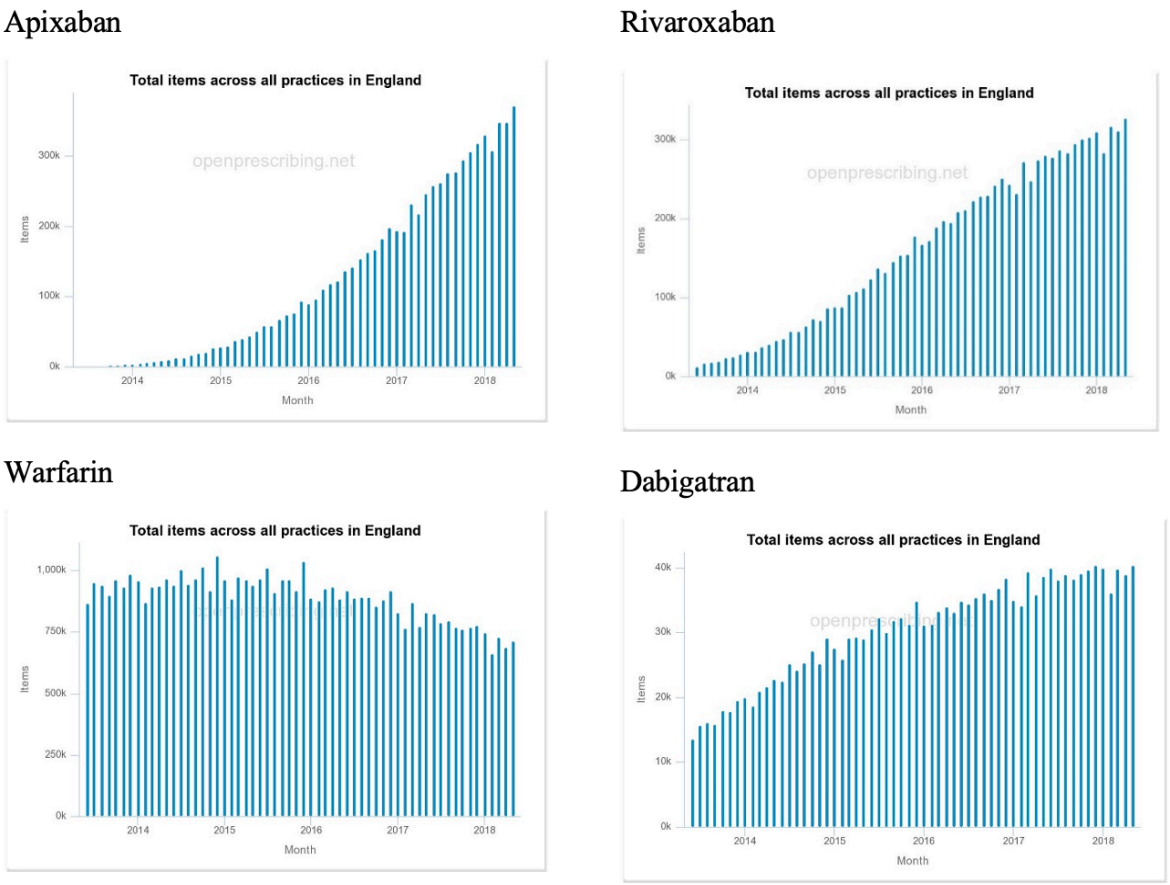


Figure 2.1 Monthly prescribing of oral anticoagulants January 2013 to May 2018
Source: OpenPrescribing.net, EBM DataLab, University of Oxford, 2018
Note: prescriptions are for all indications combined (atrial fibrillation, deep vein thrombosis, post-operative prophylactic use etc).

Having provisionally selected ARISTOTLE as the reference trial the next steps of feasibility assessment were:

i) to assess the study protocol and results in more detail and

ii) estimate the available sample size in the UK EHR by application of the trial eligibility criteria to a January 2018 extract of CPRD Gold patients prescribed apixaban.

2.2.3. Feasibility of emulating the ARISTOTLE trial

2.2.3.1. Mapping of ARISTOLE inclusion and exclusion criteria

ARISTOTLE inclusion and exclusion criteria were extracted from the trial protocol and considered in detail to determine if they were suitable for the reference trial emulation. Further detail on the application of the ARISTOTLE eligibility criteria are given in Chapter 4. Most of the criteria were deemed suitable to be applied to CPRD with Read code and medication codelists created for these criteria (codelists presented in the Appendix).

2.2.3.2. Results of feasibility analysis - application of ARISTOTLE trial criteria to CPRD Gold patients

The selected criteria were applied to a January 2018 extract of CPRD Gold patients prescribed apixaban or warfarin in the period 01 January 2013 to 31 January 2018 to determine the feasibility of emulating ARISTOTLE as a reference trial in UK EHRs. Figure 2.2 shows the number of patients excluded by each criterion. Overall, out of 13 332 patients with a prescription for apixaban and diagnosis of AF 63% (8 407) were trial-eligible, and of 68 113 patients with a prescription for warfarin in the study period and diagnosis of AF 62.3% (45 435) were trial eligible (Figure 2.2).

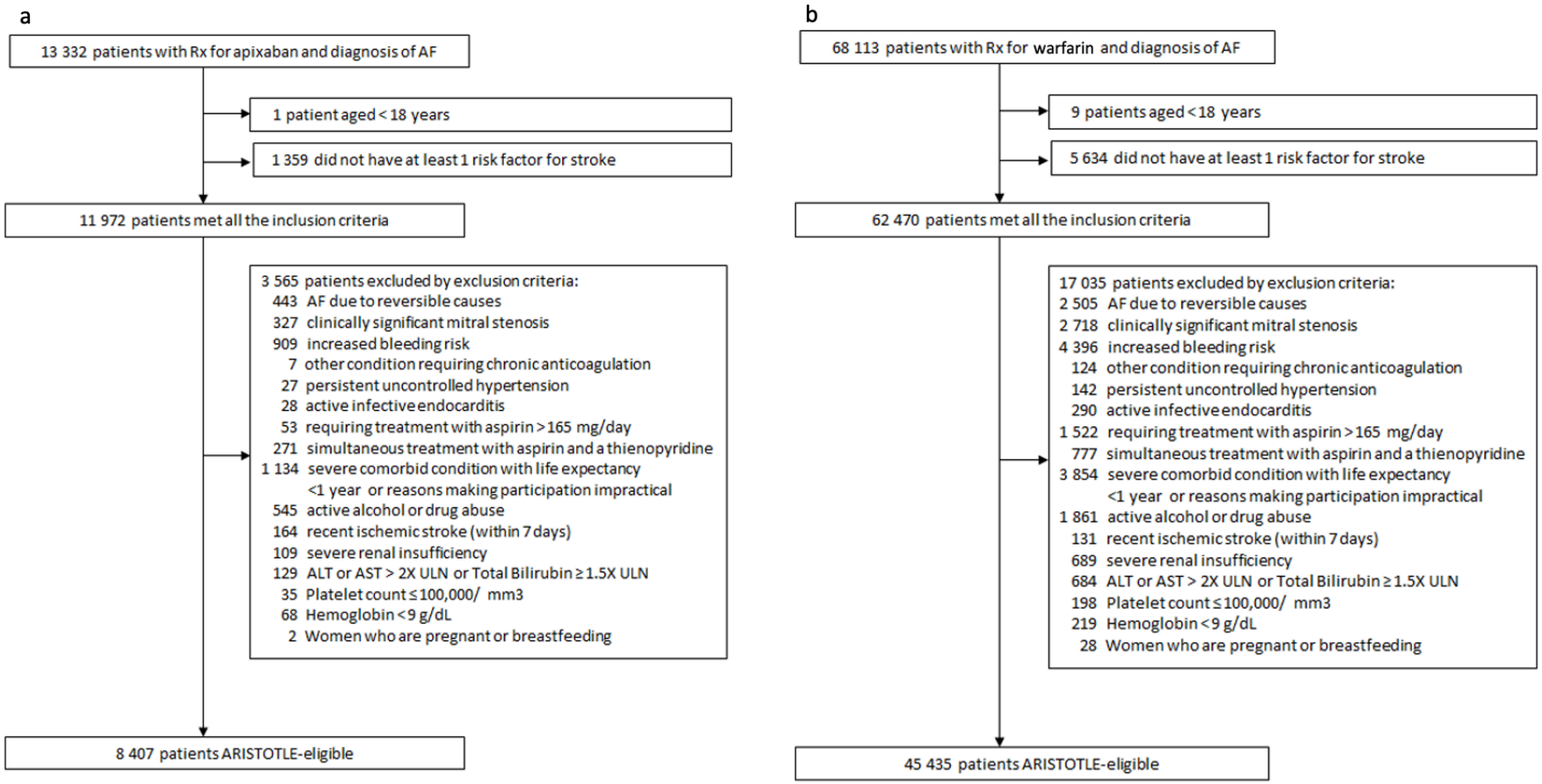


Figure 2.2 Estimate of number of CPRD Gold patients with a prescription for a) apixaban and b) warfarin meeting eligibility criteria for ARISTOTLE

Note: ARISTOTLE eligibility criteria were applied to patients with a prescription for apixaban or warfarin in the study period 01 January 2013 to 31 January 2018.

2.2.3.3. Feasibility calculations

In ARISTOTLE there were 9120 subjects in the apixaban arm therefore it was estimated a minimum of 15,000 EHR patients were needed for matching to be feasible. It was considered unlikely there would be enough patients in CPRD Gold for the project given that only ~8400 patients were eligible in the January 2018 extract. Estimates of number of patients with a prescription of apixaban and diagnosis of AF were obtained from the THIN database and CPRD Aurum providers. These numbers were combined making an allowance for duplicate patients between databases to estimate the number of unique trial-eligible EHR apixaban patients (Table 2.2). The calculations demonstrated that the objective of emulating ARISTOTLE using UK EHRs was feasible.

Table 2.2 Estimated number of unique UK EHR patients with exposure to apixaban eligible for ARISTOTLE using 2018 extract

	Patients with AF and a prescription for apixaban	Crossover with CPRD Gold (%)	Number of patients trial eligible	Number of unique trial eligible patients
CPRD Gold	13 331	NA	8 407	8 407
THIN	16 446	TBC	est. 10 000	est. 5 000
CPRD Aurum	5 318 (1 758 in Gold)	33%	est. 3 350	est. 2 200
TOTAL				est. min 15 000

est.=estimated; min=minimum; NA= not applicable; TBC=to be confirmed.
Number of patients trial eligible estimated assuming 63% of patients with AF and prescription for apixaban eligible for ARISTOTLE based on results of application of criteria to CPRD GOLD.

In order to gain an understanding of the reference trial, a background to atrial fibrillation, treatments for atrial fibrillation, and the pivotal RCTs in this area are given below.

2.3. Atrial Fibrillation

Atrial fibrillation (AF) is a heart condition that causes a patient's heart to beat abnormally. Electrical impulses control the movement of the atria (top two chambers) in the heart; in a healthy heart these impulses are regular whereas in atrial fibrillation there are periods of chaotic electrical impulses causing quivering or 'fibrillation' of the atria. Atrial fibrillation is thought to be caused by damage to cardiac tissue or electrical signalling; this damage can be

caused by ageing and/or by common conditions such as infection, high blood pressure, diabetes, and coronary heart disease.

Patients with atrial fibrillation often experience no symptoms, whereas in some cases AF can be felt by the patient as heart palpitations in which they experience a racing heart or fluttering sensation, or the patient may experience symptoms such as dizziness, breathlessness or chest pain (43).

Diagnosis of AF is typically made via 12-lead electrocardiogram (ECG) or for patients presenting with suspected paroxysmal AF, 24-hour or longer time period ambulatory (Holter) ECG monitoring may be needed to capture an episode of AF (44).

2.3.1. Epidemiology of atrial fibrillation

AF is the most common sustained cardiac arrhythmia with an estimated prevalence of 1.4 million people in England (2016) representing 2.5% of the overall population of England (45). The Framingham heart study estimated the lifetime risk of developing AF to be 26.0% (95% CI 24.0%, 27.0%) for men aged 40, and 23.0% (95% CI 21.0%, 24.0%) for women aged 40 (46).

A trend of increasing age-adjusted prevalence of atrial fibrillation was seen in the Framingham study(47) with the authors suggesting the changes may be linked to improved detection of AF and an increased prevalence of the risk factors of obesity and diabetes in more recent time periods.

2.3.1.1. Risk factors for atrial fibrillation

The most important risk factor for the development of AF is advanced age, with the analysis of the Framingham cohort by Schnabel et al(47) showing that, compared with those aged 50-59, patients aged 80-89 had 9 times the risk of AF, and those aged 70-79 had 7 times the risk

of AF. The Framingham study also showed male sex, obesity, treatment for hypertension, left ventricular hypertrophy, significant heart murmur, and myocardial infarction to be associated with higher risk of AF (Table 2.3) (47). Additional risk factors associated with AF, from studies summarised by Gahungu et al(44) and presented in Table 2.3, include valvular heart disease associated with approximately double the risk (48), obstructive sleep apnea with 2 to 3 times increased risk (49), severe chronic obstructive pulmonary disease (COPD) (50), and CKD (51).

Hypertension treatment is an important modifiable risk factor for the development of AF (52, 53). Changes to the left atrium, known as left atrial remodelling, are seen in patients with hypertension and in conditions such as heart failure, diabetes, and obesity; this remodelling appears to be "a crucial substrate for atrial fibrillation and stroke" (54). Excessive alcohol exposure appears to increase the risk of AF (55) with alcohol known to have effects on the electrical activity of the heart and to also increase the risk of hypertension.

Table 2.3 Risk factors for atrial fibrillation

Risk factor	Statistic	Association with AF
Age (years) (47)	HR (95% CI)	
50-59		1.00 (ref)
60-69		4.98 (3.49, 7.10)
70-79		7.35 (5.28, 10.2)
80-89		9.33 (6.68, 13.0)
Male vs female sex (47)	HR (95% CI)	1.49 (1.23, 1.80)
Body mass index (47)	HR (95% CI)	
Normal (<25 kg/m^2)		1.00 (ref)
Overweight (25-30 kg/m^2)		1.13 (0.87, 1.46)
Obese (≥31 kg/m^2)		1.37 (1.05, 1.78)
Hypertension treatment (47)	HR (95% CI)	1.32 (1.08, 1.60)
Diabetes (47)	HR (95% CI)	1.25 (0.98, 1.60)
Left ventricular hypertrophy on ECG (47)	HR (95% CI)	2.50 (1.21, 3.83)
Significant heart murmur (47)	HR (95% CI)	1.58 (1.09, 2.29)
Heart failure (47)	HR (95% CI)	1.43 (0.85, 2.40)
Myocardial infarction(47)	HR (95% CI)	1.46 (1.07, 1.98)
Alcohol (55)	RR (95% CI)	
Non-drinker		1.00 (ref)
>0 to 2 drinks/day		1.00 (0.92, 1.09)
>2 to 3 drinks/day		1.11 (0.98, 1.25)
>3 to 4 drinks/day		1.22 (1.02, 1.46)
>4 drinks/day		1.50 (1.22, 1.85)
Valvular heart disease (48)	RR (95% CI)	2.42 (1.62-3.60)
Obstructive sleep apnoea (49)	HR (95% CI)	
None (AHI <5)		1.00 (ref)
Mild (AHI 5-14.9)		2.12 (1.12, 2.80)
Moderate (AHI 15-29.9)		2.66 (1.98, 3.57)
Severe (AHI ≥30)		3.31 (2.53, 4.35)
COPD (50)	RR (95% CI)	
FEV_1 ≥80%		1.00 (ref)
FEV_1 60-80%		1.28 (0.79, 2.06)
FEV_1 <60%		2.53 (1.45, 4.42)
Chronic Kidney Disease (51)	OR (95% CI)	
None		1.00 (ref)
Stage 1-2		2.67 (2.04, 3.48)
Stage 3		1.68 (1.26, 2.24)
Stage 4-5		3.52 (1.73, 7.15)

AF = atrial fibrillation; AHI = apnea-hypopnea index; CI = confidence interval; COPD = chronic obstructive pulmonary disorder; ECG = electrocardiogram; HR = hazard ratio; OR = odds ratio; RR = rate ratio;
Source: Summary table presented by Gahungu et al(44) summarising results from 5 studies.
(47) Hazard ratios age-adjusted and sex-adjusted for AF with onset 1998-2007 from Framingham Heart Study [Schnabel et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study] (48) Risk ratio from Psaty et al.; (49) Hazard ratio from Cadby et al.; (50) Risk ratio from Buch et al.; (51) Odds ratio from Baber et al.

The increasing size of older age groups in countries such as the UK along with a trend of increasing obesity and diseases associated with obesity such as diabetes means the burden of

atrial fibrillation is expected to increase further over time (in the absence of any other changes).

2.3.2. Stroke risk in atrial fibrillation

The abnormal heart rhythm characteristic of AF is not in itself life-threatening, however, AF greatly increases the risk of blood clots forming in the heart which can travel to other areas of the body causing serious outcomes such as stroke and systemic embolism.

A review by Watson et al (56) provides a summary of the current understanding of the cause of the higher risk of formation of thrombi in AF. In this review the authors note that the formation of thrombi within the heart with AF is thought to be linked to the abnormal rhythm itself causing changes in blood flow leading to stasis, changes in the size of the left atrial appendage, dilation of the left atria, changes in the extracellular matrix and endothelial damage, and changes in platelets and other blood factors associated with coagulation.

Patients with AF have a five-fold increased risk of stroke compared with people without AF (57) and around a quarter of all strokes are attributed to this arrhythmia (58). In addition, increased levels of mortality, morbidity and disability with longer hospital stays are observed in stroke patients with AF compared with stroke patients without AF (59, 60).

The risk of stroke in patients with AF is also influenced by their age, sex, and other comorbidities that are known stroke risk factors such as hypertension and diabetes. Various stroke risk factor scoring systems have been in use over time; $CHADS_2$ proposed by Gage et al in 2001(61) was the stroke risk scoring system used in ARISTOTLE and is described in more detail in the methods section in chapter 4.

In 2010 a new stroke risk score, CHA_2DS_2-VASc, was derived by Lip et al (62) that had superior performance at discriminating between different levels of risk leading to its adoption

as the recommended stroke risk scoring system by the European Society of Cardiology guidelines for the management of AF in 2012a. The CHA_2DS_2-VASc scoring system is presented in Table 2.4.

Table 2.4 CHA_2DS_2-VASc stroke risk scoring system

Component		Points
C	Congestive heart failure or left ventricular systolic dysfunction	1
H	Hypertension	1
A_2	Age ≥ 75 years	2
D	Diabetes mellitus	1
S_2	Prior Stroke, TIA, or thromboembolism	2
V	Vascular disease	1
A	Age 65–74 years	1
Sc	Sex category female	1

TIA = transient ischemic attack. Hypertension defined as resting blood pressure > 140/90 mmHg on at least 2 occasions or requiring antihypertensive medication. Vascular disease includes prior myocardial infarction, peripheral artery disease, or aortic plaque.

The CHA_2DS_2-VASc score is calculated by summing the points for the different components resulting in a total score ranging from 0 to 9. The score can be used to predict the patient's annual stroke risk with a higher score corresponding to a higher stroke risk (Appendix 1). CHA_2DS_2-VASc is specified in the NICE guidance to be used to determine whether a patient should be offered anticoagulation for stroke prevention:

- Women with CHA_2DS_2-VASc score ≥ 2
- Men with CHA_2DS_2-VASc score ≥ 1

2.4. Treatment of atrial fibrillation

The NICE guidance on diagnosis and management of atrial fibrillation (published 27 April 2021 and last updated 30 June 2021) makes recommendations on the diagnosis and treatment pathway for patients with suspected atrial fibrillation in England and Wales. This pathway is described in detail in the Appendix and a brief textual summary of this pathway given below. Therapies for atrial fibrillation target two key areas – i) stroke prevention, and ii) rate and rhythm control.

2.4.1. Stroke prevention in atrial fibrillation

A key aspect of treatment for atrial fibrillation focuses on reducing the risk of stroke. When deciding whether to offer oral anticoagulant (OAC) therapy to a patient it is also necessary to consider the risk of bleeding. NICE recommends use of the ORBIT tool (63), to estimate a patient's risk of bleeding (Appendix 1).

Bleeding risk is an important consideration when deciding whether to offer a patient with AF anticoagulation therapy for stroke prophylaxis given that major bleeds in themselves are a major source of morbidity and mortality. Among the most serious major bleeding events is intracranial haemorrhage (ICH), with patients on vitamin k antagonist (VKA) oral anticoagulants found to have an absolute risk of ICH of approximate 1% per year, a risk 7- to 10- times higher compared to nonanticoagulated patients (64). Approximately 60% of intracranial haemorrhages are fatal (64).

Oral anticoagulants have been in use for the prevention of stroke in patients with atrial fibrillation since warfarin was approved by the FDA in 1954. Whereas antiplatelet drugs such as aspirin work by inhibiting platelet aggregation, the anticoagulant class of drugs have a mechanism of action that involves blocking pathways in the coagulation cascade. Multiple

pathways are involved in the coagulation cascade with different anticoagulants working by blocking different pathways as shown in Figure 2.3 from Paulus E et al (65).

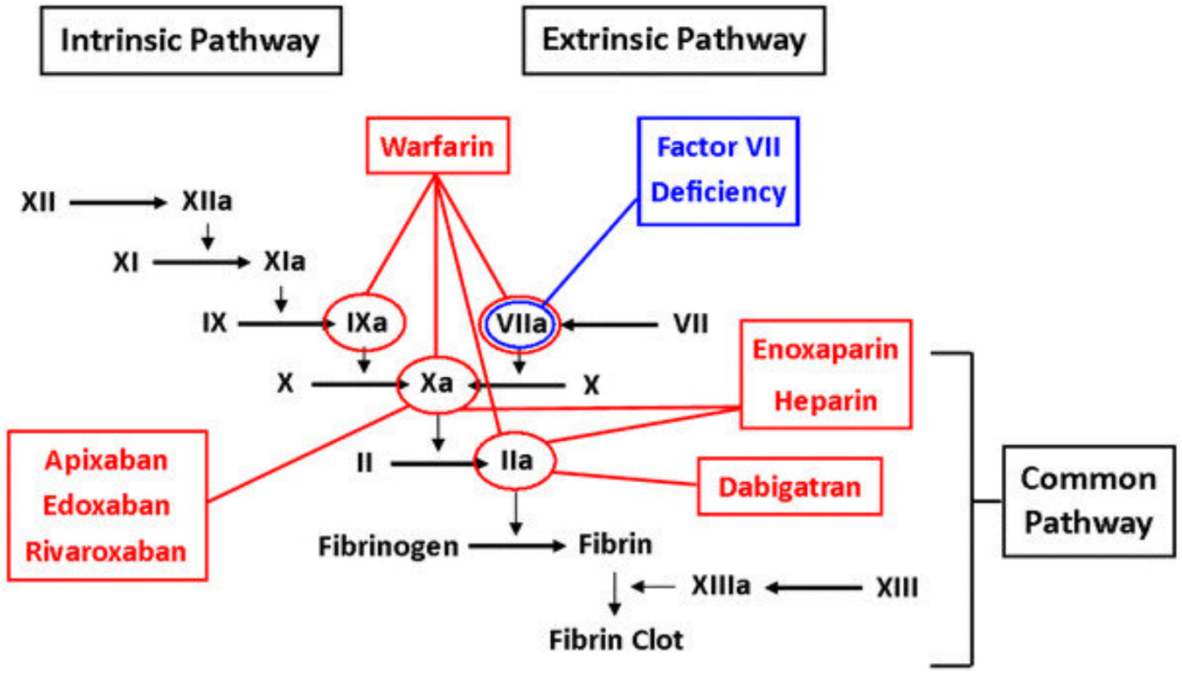


Figure 2.3 Effects of multiple anticoagulant medications on the coagulation cascade from Paulus E., et al. Anticoagulation Therapy Considerations in Factor VII Deficiency. Dec 2016 Drug Safety - Case Reports 3(1). doi:10.1007/s40800-016-0031-y. License CC BY-NC 4.0 **(65)**

2.4.1.1. Vitamin K antagonists

The class of vitamin K antagonists reduce clotting by blocking an enzyme (vitamin K epoxide reductase) involved in the reactivation of vitamin K1; vitamin K1 is required for the action of several coagulation factors (II, VII, IX, and X) (66). Warfarin is the most commonly used VKA in the UK though other vitamin K antagonists such as acenocoumarol, phenindione, and phenprocoumon are available with some more commonly used in other countries.

A meta-analysis including 29 RCTs looking at the efficacy of warfarin and antiplatelet agent by Hart et al (67) showed that compared to control (no treatment), warfarin reduces the risk of stroke by 64% (95% CI, 49% to 74%).

When warfarin therapy is initiated there is a delay of several days for full antithrombotic effects to occur following the warfarin-mediated reduction in factor II; in addition, there is temporarily an increased risk of thrombogenesis due to warfarin causing a decline in protein C (a protein with anticoagulant effects) levels and activity in the first few days post-initiation.

Warfarin and other VKA therapy is monitored regularly by measuring a patient's international normalised ratio (INR). INR is a standardised measure of how long it takes the blood to clot:

$$INR = (\frac{PT_i}{PT_C})^{ISI}$$

Where
PT_i = the prothrombin time (seconds) of the patient
PT_c = the prothrombin time (seconds) of a standard sample used in the laboratory analysing the patient's sample
ISI = International Sensitivity Index, a measure of the sensitivity of the tissue factor used to analyse PT_i and PT_c typically in the range of 0.9-1.7.

Therapy with vitamin k antagonists typically aims to raise a patient's INR to fall within the therapeutic range [2, 3] in which the risk of both ischemic and bleeding events are minimised. As the INR falls below 2 the risk of ischemic events increases and at values above 3 the risk of bleeding rises.

On initiation of VKA therapy, INR is checked daily until in therapeutic range, then 3 times weekly for 2 weeks, then less often, according to the stability of the results. A measure to estimate a patient's INR control over time is the time in therapeutic range (TTR); different methods for estimation of TTR exist with 2 of the most commonly used being: 1) a simple proportion of INR readings within the optimal range for a given time frame and 2) Rosendaal's method using linear interpolation between INR readings to estimate the proportion of days a patient's INR has been within optimal range.

Warfarin and the other VKAs were renowned for having many treatment interactions and needing frequent monitoring and dose adjustments to stay within the therapeutic range of anticoagulant action as summarised by Hirsh et al(68): Warfarin has interactions with a wide range of drugs such as metronidazole which inhibits warfarin clearance, barbiturates and carbamazepine which increase hepatic warfarin, and aspirin which increases the risk of bleeding. Diet also interacts with response to warfarin with increased intake of vitamin K (present in green vegetables) leading to a reduction in the anticoagulant response to warfarin. Genetics influences the warfarin dose-response relationship most notably in common mutations in coding for cytochrome P450.

Given the challenge in maintaining INR in therapeutic range and the complex safety profile of warfarin it was hoped that the introduction of DOACs would provide a safer and easier to manage long term anticoagulation therapy for AF patients.

2.4.1.2. Direct-acting oral anticoagulants

The direct-acting oral anticoagulants (DOACs) consist of 5 drugs:

- Dabigatran, approved in the UK in 2008
- Rivaroxaban, approved in the UK in 2008
- Apixaban, approved in the UK in 2011
- Edoxaban, approved in the UK in 2015

The DOACs are indicated for a number of conditions including the prevention and treatment of deep vein thrombosis, pulmonary embolism, and prevention of stroke in Non-valvular atrial fibrillation (NVAF). DOACs have many advantages over VKA, most noticeably that they do not require frequent monitoring of INR and personalised dose-adjustments. The DOACs also benefit from a faster onset of action and fewer food and drug interactions compared with VKA (69).

2.4.1.3. RCTs in Direct-acting oral anticoagulants in AF

Results from the pivotal RCTs that informed the NICE guidance on oral anticoagulation in AF are summarised in Table 2.5.

Table 2.5 Summary of RCTs in DOACs vs warfarin for stroke prevention in non-valvular atrial fibrillation

Trial	Details
RE-LY (dabigatran vs warfarin)(70) Patients enrolled: Dec 2005 – Dec 2007 LPLV: Mar 2009 Results published NEJM Sep 17 2009	NI margin: 1.46 with one-sided alpha level of 0.025. Inclusion: AF and (previous stroke or TIA or LVEF < 40% in last 6 months, symptomatic heart failure, age ≥ 75 years or age 65 to 74 years plus [diabetes mellitus or hypertension, or coronary artery disease]) Key exclusion: severe heart-valve disorder, stroke within 14 days or severe stroke within 6 months, increased risk of haemorrhage, creatinine clearance < 30 ml/min, active liver disease Blinded independent central adjudication of outcomes. Patients recruited from 951 sites in 44 countries. Randomised to dabigatran 150 mg twice daily, dabigatran 110 mg twice daily, or adjusted-dose warfarin. Dabigatran blinded, warfarin open-label. 18,113 patients: 6076 high-dose dabigatran, 6015 low-dose dabigatran, 6022 warfarin. Median age 72, 36.4% female, mean $CHADS_2$ 2.1, 49.6% prior VKA use Median duration of follow-up 2.0 years. TTR in warfarin arm mean 64%. Results for high dose dabigatran vs warfarin, event rates and HR (CI) *Stroke/SE ITT 1.11%/yr 1.69%/yr 0.66 (0.53, 0.82) Ischemic stroke 0.92%/yr 1.20%/yr 0.76 (0.56, 0.89) Haemorrhagic stroke 0.10%/yr 0.38%/yr 0.26 (0.14, 0.49) All-cause death 3.64%/yr 4.13%/yr 0.88 (0.77, 1.00) Major bleeding 3.11%/yr 3.36%/yr 0.93 (0.81, 1.07) ICH 0.30 0.74 0.40 (0.27, 0.60)
ROCKET-AF (rivaroxaban vs warfarin)(71) Patients enrolled: Dec 2006 – Jun 2009 LPLV: May 2010 Results published NEJM Sep 8 2011	NI margin: 1.46 with a one-sided alpha level of 0.025. Inclusion: nonvalvular AF and $CHADS_2$ score ≥ 2. (history of stroke/TIA/SE or 2+ of the following risk factors: heart failure or a LVEF of 35% or less, hypertension, age ≥ 75 years, diabetes mellitus Only 10% per region allowed to have [no prior stroke/SE/TIA or up to 2 other risk factors], remainder to have prior stroke/SE/TIA or ≥ 3 risk factors 14,264 patients: 7131 rivaroxaban, 7133 warfarin. Median age 73, 39.7% female, mean $CHADS_2$ 3.5, 62.4% prior VKA use Median duration of follow-up 1.9 years.

Trial	Details
	TTR in warfarin arm mean 55%. Results for rivaroxaban vs warfarin, event rates and HR (CI) *Stroke/SE PP 1.7%/yr 2.2%/yr 0.79 (0.66, 0.96) NI, +superior Stroke/SE ITT 2.1%/yr 2.4%/yr 0.88 (0.75, 1.03) NI Stroke/SE on-treatment 1.7%/yr 2.2%/yr 0.79 (0.66, 0.96) Ischemic stroke 2.11%/yr 2.27%/yr 0.94 (0.75, 1.17) Haemorrhagic stroke 0.41%/yr 0.71%/yr 0.59 (0.37, 0.93) All-cause death PP 1.9%/yr 2.2%/yr 0.85 (0.70, 1.02) All-cause death ITT 4.5%/yr 4.9%/yr 0.92 (0.82, 1.03) Major bleeding 5.6%/yr 5.4%/yr 1.04 (0.90, 1.20) ICH 0.5 0.7 0.67 (0.47, 0.93)
ARISTOTLE (apixaban vs warfarin)(42) Patients enrolled from Dec 2006 – Apr 2010 Results published NEJM Sep 15 2011	NI margin for 2-sided CIs: upper 99% CI < 1.44, upper 95% CI < 1.38 (depending on regulator) Inclusion: AF and at least one of risk factor for stroke (age of at least 75 years; previous stroke, TIA, or SE; symptomatic heart failure within previous 3 months or LVEF <=40%; diabetes mellitus; or hypertension requiring pharmacologic treatment. Key exclusion criteria: AF due to a reversible cause, moderate or severe mitral stenosis, conditions other than AF that required anticoagulation (e.g. prosthetic heart valve), stroke within previous 7 days, concomitant aspirin >165 mg a day or both aspirin and clopidogrel, severe renal insufficiency (serum creatinine >2.5 mg/dL or creatinine clearance <25 m/min) Recruited 18,201 patients at 1034 clinical sites in 39 countries 18,201 patients: 9120 apixaban, 9081 warfarin. Median age 70, 35.3% female, mean $CHADS_2$ 2.1, 57.1% prior VKA use Median duration of follow-up 1.8 years. TTR in warfarin arm mean 62.2%, median 66% Results for apixaban vs warfarin, event rates and HR (95% CI) *Stroke/SE ITT 1.27%/yr 1.60%/yr 079 (0.66, 0.95) Ischemic stroke 0.97%/yr 1.05%/yr 0.92 (0.74, 1.13) Haemorrhagic stroke 0.24%/yr 0.47%/yr 0.51 (0.35, 0.75) All-cause death 3.52%/yr 3.94%/yr 0.89 (0.80, 0.998) Major bleeding 2.13%/yr 3.09%/yr 0.69 (0.60, 0.80)
ENGAGE AF-TIMI 48 (edoxaban vs warfarin) (72) Patients enrolled from Nov 2008 to Nov 2010 Results published NEJM Nov 28 2013	NI margin: upper boundary of one-sided 97.5% CI <1.38. Inclusion: AF, $CHADS_2$ score ≥ 2, age ≥21 years. Key exclusion: reversible cause of AF, creatinine clearance < 30 ml /min; high bleeding risk; use of dual antiplatelet therapy; moderate-to-severe mitral stenosis; other indications for anticoagulation; acute coronary syndromes, coronary revascularization, stroke within 30 days prior; inability to adhere to study procedures Randomised to warfarin, 60mg or 30mg edoxaban (edoxaban dose halved in both arms if creatinine clearance of 30-50 ml/min, weight <=60 kg, or concomitant use of verapamil, quinidine, or dronedarone).

Trial	Details
	Randomisation stratified by: $CHADS_2$ score [2 or 3] vs [4, 5, or 6], and need for reduced edoxaban dose. 21,105 patients: 7035 high-dose edoxaban, 7034 low-dose edoxaban, 7036 warfarin. Median age 72, 38.1% female, mean $CHADS_2$ 2.8, prior VKA use 58.9% Median duration of follow-up 2.8 years TTR in warfarin arm median 68.4% (IQR, 56.5 to 77.4), mean (±SD) 64.9±18.7% Results for high-dose edoxaban vs warfarin, event rates (%/year) and HR (95% CI) *Stroke/SE mITT 1.18%/yr 1.50%/yr 0.79 (0.63, 0.99) Stroke/SE ITT 1.57%/yr 1.80%/yr 0.87 (0.73, 1.04) Stroke/SE on-treatment 1.18%/yr 1.50%/yr 0.79 (0.63, 0.99) Ischemic stroke 1.25%/yr 1.25%/yr 1.00 (0.83, 1.19) Haemorrhagic stroke 0.26%/yr 0.47%/yr 0.54 (0.38, 0.77) All-cause death 3.99%/yr 4.35%/yr 0.92 (0.83, 1.01) Major bleeding 2.75%/yr 3.43%/yr 0.80 (0.71, 0.91)

AF=atrial fibrillation; IQR=interquartile range; ITT=intent-to-treat; LVEF=left ventricular ejection fraction; mITT=modified intent-to-treat, in ENGAGE AF-TIMI 48 this was randomised patients that received at least one dose of study drug; SE=systemic embolism; TIA=transient ischemic attack; TTR=time in therapeutic range

The design of RE-LY was criticised, specifically the lack of blinding in the warfarin arm was thought to have caused 'differential treatment of patients during the study period' (FDA review (73)). The FDA review also noted a trend towards increased mortality in the patients on warfarin in sites with inferior INR control whereas in sites with TTR ≥ 67% there was increased risk of death in dabigatran compared with warfarin(73). The high rate of intracranial haemorrhage observed in the warfarin arm (with an estimated rate of 0.76% per year) when compared with the warfarin arms in other RCTs in AF (with 0.3% and 0.45% from Cochrane reviews) was also a cause for concern (74). The high rate of concomitant treatment with antiplatelets was criticised given the approximately doubling of major bleeding events in patients taking antiplatelets with anticoagulants (74).

ROCKET-AF was criticised for the low TTR reported in the warfarin arm of the trial (55%) and ambiguity over the method used to calculate TTR (75). Furthermore, the BMJ uncovered that defective point of care devices for patient monitoring of INR were used in ROCKET-AF

(76): the BMJ hypothesised this fault could have led to patients having their warfarin dose increased unnecessarily leading to a greater risk of bleeding (76).

All of the DOAC RCTs for the indication of prevention of stroke and systemic embolism in AF had similar study designs with key differences being that only RE-LY used open-label warfarin (all other RCTs used dummy INR testing and kept both treatment arms blinded), and minor differences in the eligibility criteria. Comparison of the baseline characteristics showed all RCTs had similar median age (range 70 to 73), proportion of female participants (range 35.3% to 39.7%), however differences in baseline stroke risk were seen, with mean CHADS score 2.1, 3.5, 2.1, and 2.8 in RE-LY, ROCKET-AF, ARISTOTLE and ENGAGE AF-TIMI respectively. The proportion of participants with prior VKA exposure ranged from 49.6% to 62.4%. A key difference between the trials was the quality of INR control in the warfarin arm with a lower time in therapeutic range in ROCKET-AF (55%) than the other RCTs (RE-LY 64%, ARISTOTLE 62.2%, and ENGAGE AF-TIMI 64.9%).

In terms of results, all RCTs met their criteria for non-inferiority and were successful in subsequent testing for superiority against warfarin for the primary endpoint of stroke or systemic embolism. Point estimates for the hazard ratios indicated approximately 20% lower risk vs warfarin for rivaroxaban, apixaban, and edoxaban, and a 34% lower risk for dabigatran. The main driver of the lower risk in the DOACs vs warfarin was the lower risk of haemorrhagic stroke compared with warfarin in all the trials whereas the rate of ischemic stroke was mostly similar between the DOACs and warfarin (except for dabigatran which showed a lower risk in the high-dose dabigatran arm).

Only apixaban showed a statistically significant lower risk of all-cause death when compared with warfarin. Safety results showed the DOACs to be non-inferior for major bleeding vs warfarin for dabigatran and rivaroxaban, and superior for major bleeding vs warfarin for

apixaban and edoxaban. This difference in mortality and safety results is likely to explain the observed trend of greater use of apixaban in real world patients in the UK compared with the other DOACs. Furthermore, despite dabigatran having been marketed as not requiring monitoring, there was evidence from an investigation by the BMJ in 2014 (77) of substantial variability in serum levels of dabigatran particularly in the elderly meaning monitoring of drug levels and dose adjustment may be advised.

2.4.1.4. Choice of oral anticoagulant in AF

The current first-line treatment option per the NICE guidance consists of the DOACs); NICE recommends the patient and clinician discuss the risks and benefits of the different DOACs available and select the DOAC most suitable to the patient. Each DOAC has a slightly different pharmacokinetic/pharmacodynamic and safety profile which means depending on patient factors such as renal function one DOAC may be preferred over another.

For patients that have contraindications to, or cannot tolerate the DOACs, a vitamin K antagonist such as warfarin can be offered as an alternative oral anticoagulant. The key contraindications to DOACs are valvular AF (moderate or severe mitral stenosis), presence of a mechanical heart valve, and antiphospholipid syndrome. Warfarin is also the only OAC not contra-indicated in severe renal impairment (creatinine clearance <15 mL/min) whereas for patients with moderate renal impairment (15-29 mL/min) there is a choice of suitably adjusted dose of apixaban, edoxaban, or rivaroxaban. Apixaban tends to be the favoured DOAC for patients with renal impairment given it has the lowest rate of renal elimination of all DOACs (27%) (78).

The anticoagulation of patients with advanced CKD poses a particular challenge as DOACs are contra-indicated and VKA therapy is more likely to result in out of range INR with the associated risks in these patients; a study in patients with AF on dialysis found those treated

with warfarin had a 44% higher risk of bleeding with no benefit in risk of stroke when compared with patients not on warfarin (adjusted HR 1.14, 95% CI 0.78, 1.67) (79).

Patients that initiated VKA therapy prior to the availability of the DOACs may prefer to stay on the VKA for reasons of familiarity and stability; NICE recommends these patients should be offered the option to switch to a DOAC and advised on the relative risks and benefits of such a switch. Switching from a VKA to a DOAC is recommended in cases where a patient meets any of the following criteria as indicating poor quality VKA therapy:

- Two INR values higher than 5, or one INR value higher than 8 within the past 6 months.
- Two INR values less than 1.5 within the past 6 months.
- Time in therapeutic range (TTR) less than 65%.

2.4.1.5. Trends in OAC use

A study by Afzal et al (80) looked at prescribing trends in OACs (for all indications) in the primary care setting in England in the period 2009–2019; Afzal found that the use of DOACs as a proportion of total OAC prescriptions increased from 16% in 2015 to 62% in 2019. The Afzal study also reported estimates for the proportion of total OAC prescriptions (%) for each individual DOAC and warfarin, showing apixaban to be the most commonly prescribed DOAC at 31.8%, followed by rivaroxaban (23.4%), edoxaban (4.1%), and dabigatran (2.5%). Warfarin represented 38.1% of all anticoagulant prescriptions in the study in 2019, a marked decrease from 2015 when it represented 83.6%. Afzal also noted a further increase in apixaban prescriptions in 2020 with apixaban overtaking warfarin.

2.5. ARISTOTLE

ARISTOTLE was the key pivotal trial for the DOAC apixaban for the indication of prevention of stroke and systemic embolism in patients with non-valvular atrial fibrillation and at least one risk factor for stroke. ARISTOTLE was a large multi-country study which

enrolled participants from December 2006 through to April 2010. ARISTOTLE enrolled participants from 39 countries across 4 geographical regions (North America, South America, Europe, Asian Pacific) with a total of 1034 different sites (for example a hospital or clinic) involved.

The primary objective was to determine whether apixaban was noninferior to warfarin in preventing stroke or systemic embolism in patients with atrial fibrillation and at least one additional risk factor for stroke. The primary safety outcome of the trial was major bleeding. The key secondary objectives included assessing whether apixaban was superior to warfarin for: the primary outcome, key safety outcome, and death from any cause.

The key inclusion and exclusion criteria are listed in Table 2.5. ARISTOTLE included both patients with no prior VKA exposure and patients currently or previously exposed to VKA with randomisation stratified by prior VKA exposure status (naïve or experienced).

2.5.1. ARISTOTLE Methods

ARISTOTLE analysed outcome events using Cox proportional hazards models stratified by geographical region and prior VKA status (experienced, naïve). A hierarchical testing approach was used in which first the primary efficacy outcome (time to stroke or systemic embolism) was tested for non-inferiority, if the non-inferiority criteria were met this was to be followed by testing the primary outcome and key secondary outcomes for superiority in order. Safety outcomes were analysed in a similar way testing for the equality of the rates.

2.5.2. ARISTOTLE Results

A total of 18,201 participants were randomised in a 1:1 ratio, with 9120 participants assigned to the apixaban group and 9081 to the warfarin group. The two groups were well balanced with respect to baseline characteristics (Table 2.6) with a median age of 70, 35% participants were female and a mean $CHADS_2$ score of 2.1.

Table 2.6 Baseline table for NEJM published results of ARISTOTLE

Characteristic - n(%) unless specified	Apixaban (N = 9 120)	Warfarin (N = 9 081)
Age – years , median (IQR)	70 (63, 76)	70 (63, 76)
Female sex	3 234 (35.5)	3 182 (35.0)
Systolic blood pressure – mmHg, median (IQR)	130 (120, 140)	130 (120, 140)
Weight – kg, median (IQR)	82 (70, 96)	82 (70, 95)
Prior myocardial infarction	1319(14.5)	1266 (13.9)
Prior clinically relevant or spontaneous bleeding	1525 (16.7)	1515 (16.7)
History of fall within previous year	386 (4.2)	367 (4.0)
Prior use VKA >30 days	5 208 (57.1)	5 193 (57.2)
Qualifying risk factors		
Age ≥ 75 years	2 850 (31.2)	2 828 (31.1)
Prior stroke, TIA, or SE	1 748 (19.2)	1 790 (19.7)
Heart failure or reduced LVEF	3 235 (35.5)	3 216 (35.4)
Diabetes	2 284 (25.0)	2 263 (24.9)
Hypertension req. treatment	7 962 (87.3)	7 954 (87.6)
$CHADS_2$ score, mean ± SD	2.1 ± 1.1	2.1 ± 1.1
$CHADS_2$ =0	54 (0.6)	58 (0.6)
$CHADS_2$ =1	3 046 (33.4)	3 025 (33.3)
$CHADS_2$ =2	3 262 (35.8)	3 254 (35.8)
$CHADS_2$ ≥3	2 758 (30.2)	2 744 (30.2)
Medications at index date		
ACE inhibitor or ARB	6 464 (70.9)	6 368 (70.1)
Amiodarone	1 009 (11.1)	1 042 (11.5)
Beta-blocker	5 797 (63.6)	5 685 (62.6)
Aspirin	2 859 (31.3)	2 773 (30.5)
Clopidogrel	170 (1.9)	168 (1.9)
Digoxin	2 916 (32.0)	2 912 (32.1)
Calcium channel blocker	2 744 (30.1)	2 823 (31.1)
Statin	4 104 (45.0)	4 095 (45.1)
Non-steroidal anti-inflammatory	752 (8.2)	768 (8.5)
Gastric antacid drugs	1 683 (18.5)	1 667 (18.4)
Renal function, creatinine clearance		
Normal, >80 ml/min	3 761 (41.2)	3 757 (41.4)
Mild imp., >50 to 80 ml/min	3 817 (41.9)	3 770 (41.5)
Moderate imp. (>30 to 50 ml/min)	1 365 (15.0)	1 382 (15.2)
Severe imp. (≤30 ml/min)	137 (1.5)	133 (1.5)
Not reported	40 (0.4)	39 (0.4)
Ethnicity		
White	7 536 (82.6)	7 493 (82.5)
Black	125 (1.4)	102 (1.1)
Asian	1 310 (14.4)	1 332 (14.7)
Other	149 (1.6)	153 (1.7)

ACE = angiotensin-converting enzyme; ARB = angiotensin-receptor blocker; IQR=interquartile range; LVEF= left ventricular ejection fraction; SD = standard deviation; TIA = transient ischemic attack

Source: C B. Granger et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2011; 365:981-992, doi: 10.1056/NEJMoa1107039

Median duration of follow-up was 1.8 years. ARISTOTLE successfully demonstrated non-inferiority of apixaban vs warfarin for the primary endpoint and showed superiority (Table 2.7). Apixaban also showed superiority for all-cause death (Table 2.7) and major bleeding (Table 2.8).

Table 2.7 Efficacy outcomes results from ARISTOTLE

	Apixaban Group (N=9,120)		**Warfarin Group (N=9,081)**		
Outcome	Patients with Event *no.*	Event Rate *%/yr*	Patients with Event *no.*	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	212	1.27	265	1.60	0.79 (0.66,0.95)
Stroke	199	1.19	250	1.51	0.79 (0.65,0.95)
Ischemic or uncertain type of stroke	162	0.97	175	1.05	0.92 (0.74,1.13)
Haemorrhagic stroke	40	0.24	78	0.47	0.51 (0.35,0.75)
Systemic embolism	15	0.09	17	0.10	0.87 (0.44,1.75)
Key secondary outcome: death from any cause	603	3.52	669	3.94	0.89 (0.80,0.998)
Other secondary outcomes					
Stroke, systemic embolism, or death from any cause	752	4.49	837	5.04	0.89 (0.81,0.98)
Myocardial infarction	90	0.53	102	0.61	0.88 (0.66,1.17)
Stroke, systemic embolism, myocardial infarction, or death from any cause	810	4.85	906	5.49	0.88 (0.80,0.97)
Pulmonary embolism or deep-vein thrombosis	7	0.04	9	0.05	0.78 (0.29,2.10)

ref: C B. Granger et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2011; 365:981-992, doi: 10.1056/NEJMoa1107039

Table 2.8 Bleeding outcomes and net clinical outcomes results from ARISTOTLE RCT

ARISTOTLE RCT	**Apixaban Group (N=9,088)**		**Warfarin Group (N=9,052)**		
Outcome	Patients with Event *no.*	Event Rate *%/yr*	Patients with Event *no.*	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary safety outcome: ISTH major bleeding	327	2.13	462	3.09	0.69 (0.60,0.80)
Intracranial	52	0.33	122	0.80	0.42 (0.30,0.58)
Other location	275	1.79	340	2.27	0.79 (0.68,0.93)
Gastrointestinal	105	0.76	119	0.86	0.89 (0.70,1.15)
Net clinical outcomes					
Stroke, SE, or major bleeding	521	3.17	666	4.11	0.77 (0.69,0.86)
Stroke, SE, major bleeding, or death from any cause	1009	6.13	1168	7.20	0.85 (0.78,0.92)

ISTH=International Society on Thrombosis and Haemostasis; SE=systemic embolism.
ref: C B. Granger et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2011; 365:981-992, doi: 10.1056/NEJMoa1107039

2.5.3. Discussion of ARISTOTLE Results

ARISTOTLE was a well-designed RCT that demonstrated superior efficacy and safety of apixaban compared with warfarin. The results of ARISTOTLE led to apixaban being added as a recommended treatment option in the UK for patients with atrial fibrillation and additional stroke risk factor(s). The primary outcome included both ischemic and haemorrhagic strokes with the ARISTOTLE results suggesting the main driver of the overall lower rate of stroke or systemic embolism was the lower haemorrhagic stroke rate in the apixaban users (0.24%/year vs 0.47%/year for warfarin) whereas little difference was seen in the rates of ischemic stroke (0.97%/year vs 1.05%/year).

The key limitations for ARISTOTLE relate to questions on the generalisability of the results (in common with most RCTs, how applicable are the results to the target population?) and questions on the quality of the warfarin treatment in the warfarin arm given the known association between INR control in warfarin users and risk of ischemic and bleeding events. Generally, one might assume that the standard of care offered in an RCT should be delivered to an equivalent or higher standard than what is seen in routine care given the increased monitoring of patients involved in an RCT. The key measure of the standard of care for warfarin users is TTR with thresholds of 65% used by NICE to indicate acceptable control, and higher thresholds such as 70% or 75% considered for good control. The mean TTR in the ARISTOTLE warfarin arm was 62.2% (median 66%); the Appraisal Committee considering whether to recommend apixaban in the UK considered this TTR to reflect ‘what is generally seen in the UK, not what is observed in centres achieving the best time in therapeutic range, and that centres should aim for a time in therapeutic range for each individual of 70% and above’. In the NICE review of ARISTOTLE several professional groups noted that the TTR of warfarin users in ARISTOTLE may be lower than what is typical in UK clinical practice [9] , questioned whether “apixaban compared with well-controlled warfarin (TTR 75% or

more) may not be superior in the long term", and noted a trend in the analysis by TTR performed in ARISTOTLE in which event rates in the apixaban arm varied by TTR group suggesting factors other than INR control may have contributed to the results grouped by cTTR.

An analysis of the EU patient subgroup from the trial (22.3% of the participants) in the EMA commentary of the results (81) noted reduced efficacy in the patients in the EU (Stroke/SE HR 0.92 [95% CI 0.56, 1.52], all cause death HR 0.89 [95% CI 0.68, 1.18]) and noted this could be attributed to the superior INR control in the EU patients (median 0.6893) (81). The EMA commentary noted superiority for the primary endpoint was lost in sites with TTR ≥ median TTR, for all cause death in the highest TTRc quartile (>72.2%) the HR was 1.04 (95% CI 0.82, 1.33) concluding superiority was not shown in patients well controlled on VKA.

The comments in the EMA and NICE reviews suggested INR control as measured by TTR would likely be a key factor to consider in the emulation of ARISTOTLE using UK EHRs and that questions regarding the TTR seen in UK EHR compared with the RCT and relationship between the relative effectiveness and the TTR in the warfarin group could be explored as part of the ARISTOTLE emulation.

The generalisability of the ARISTOTLE results to a given population is likely to depend not only on the typical quality of warfarin therapy in that population, but also other factors relating to the similarity of the population to the ARISTOTLE participants. The Appraisal Committee noted that compared with the ARISTOTLE participants 'people treated for atrial fibrillation tended to be older and more likely to be on non-steroidal anti-inflammatory drugs (NSAIDS), which can impact on bleeding', however overall the Committee concluded that ARISTOTLE was 'broadly generalisable to the UK population'. This is an area that can be

explored during the emulation of ARISTOTLE in UK EHR – that is, how similar are the ARISTOTLE participants to UK patients with AF prescribed apixaban or warfarin? How might differences in the baseline characteristics impact the relative benefits of apixaban vs warfarin? Further, should emulation of ARISTOTLE be successful then apixaban could be compared to warfarin in patient groups excluded from or under-represented in ARISTOTLE such as the elderly, or those at increased bleeding risk.

2.6. Summary

This chapter summarised background information for the thesis. First a description of the concepts of reference trial emulation and the related target trial emulation were given. The high-level results and conclusions of the authors of the RCT-DUPLICATE initiative which has emulated a large number of RCTs using US insurance claims data were described along with the additional reference trial emulation studies identified by Baptiste (21). It was noted that both the FDA and RCT-DUPLICATE authors recommend more reference trial emulation studies are required to deepen the understanding of differences in results and the influence of the data sources and methods on the results. The objective of this thesis to attempt to emulate a reference trial using UK EHR data including matching to the trial participants and exploring the inclusion of prevalent users could therefore provide evidence on outstanding questions in the topic of reference trial emulation.

This chapter then introduced the intended data sources of this thesis consisting of UK electronic healthcare records from primary care linked to hospital admissions and mortality data, and patient level data from the target reference trial. The feasibility criteria for selecting a suitable reference trial to emulate using this data were given along with discussion of the suitable therapeutic areas based on the characteristics of typical RCT designs, outcomes, and the quality and breadth of data available in the UK EHRs. The limitations of the data and considerations of missing data in the context of longitudinal outcomes in real world data led

to the identification of suitable therapeutic areas and the selection of ARISTOTLE, which compared apixaban to warfarin for stroke prevention in patients with AF, as a potential candidate reference trial to emulate. The feasibility work conducted in which the eligibility criteria of ARISTOTLE were applied to CPRD Gold to determine whether it would be possible to emulate this reference trial was presented. The calculations showed there would be a large enough sample size available in the UK EHR data sources to emulate ARISTOTLE.

In order to understand the ARISTOTLE trial this chapter concluded with a brief summary of ARISTOTLE, a background to atrial fibrillation and the current treatment landscape for atrial fibrillation.

The next chapter will present the results of a literature review looking at real world evidence comparing apixaban to warfarin.

Chapter 3 Literature review: a scoping review on the "real world" effectiveness and/or safety of apixaban compared with VKA in stroke prevention in AF patients

This chapter aims to provide an overview of the existing literature on studies assessing the real-world effectiveness of apixaban compared with warfarin for stroke prevention in patients with AF. This chapter will present:

- An introduction and the aims of the literature review,
- The methods used to identify studies,
- Results of the literature review looking at non-interventional studies assessing effectiveness and/or safety of apixaban.
- A summary of the data sources, key methods, and results of the studies identified in the literature review.
- A conclusion of the results of the literature review.

3.1. Introduction and aims

The aim of the review is to provide a scoping review of the literature on non-interventional studies assessing effectiveness and/or safety of apixaban compared with warfarin, using systematic searches of databases. More specifically to summarise the:

1. Populations studied and how this compared to the ARISTOTLE participants

2. Design, methods and analysis approach used to estimate treatment effects and how confounding and bias were accounted for

3. Findings from the non-interventional research relative to ARISTOTLE

As this review was focused on the methods and study design rather than the result, a meta-analysis was not performed.

3.2. Methods

3.2.1. Databases and sources

Two large medical journal databases, PUBMED (1946 to present) and EMBASE (1947 to present) were searched electronically for non-interventional studies into atrial fibrillation patients treated with a VKA or a DOAC where the effectiveness and/or safety of the treatments was compared. To uncover potential "grey literature" the websites of electronic health record databases were searched for publications. The reference list of selected publications and any review articles found were also searched for relevant articles.

3.2.2. Search keywords and terms

In designing the search strategy, the research question was divided into three main categories of terms: (1) terms related to the treatments of interest (VKAs and DAOCs) (2) terms related to atrial fibrillation as the population of patients of interest and (3) terms relating to non-interventional studies. Terms within each category were combined by "or" statements while the two categories were combined by an "and" statement.

Within Pubmed which is indexed by MeSH terms, the following MeSH keywords were identified and used:

(((Dabigatran[MeSH Terms]) OR ("factor xa inhibitors"[MeSH Terms]) OR (apixaban) OR (warfarin) OR (4-Hydroxycoumarins[MeSH Terms])) AND (Observational Study[Publication Type]) AND (atrial fibrillation[MeSH Terms]))

Within EMBASE which is indexed using Emtree preferred terms, the following EMBASE terms were identified and used:

(blood clotting factor 10a inhibitor OR coumarin anticoagulant OR dabigatran OR dabigatran etexilate) AND atrial fibrillation AND observational study

3.2.3. Procedure

Titles and abstracts of all retrieved articles were screened for an initial assessment of eligibility for inclusion. For the articles that fulfilled the inclusion criteria (or articles for which eligibility was still unclear), the full text was obtained for review. Eligibility decisions were finalised based on the full-text review as necessary, and the following information was extracted from each included article: first author and year of publication, database(s) used, study design, outcome(s), exposure(s), age range criteria for selection of study population, inclusion criteria used, main study results, analysis methods used, sensitivity analyses performed.

3.2.4. Inclusion/exclusion criteria

Inclusion criteria

1. Any study comparing effectiveness and/or safety of apixaban to a vitamin K antagonist (VKA).

2. Outcome (primary or otherwise) of the study is stroke, systemic embolism, myocardial infarction (MI), bleeding, or death (alone or in combination, including synonyms of all outcomes and subtypes of outcomes such as intracranial haemorrhage).

3. Study using non-interventional data.

Exclusion criteria

Studies focused on:

1. Anticoagulation before, during, or after surgical procedures.
2. Patterns of treatment or epidemiology of AF.
3. Adherence, persistence, or interactions between treatment and factors such as diet and genetics.
4. Special patient populations such as liver failure or end stage kidney disease.
5. Only pooled DOACS or DOACs other than apixaban

6. Indications other than AF

3.2.5. Search date

Searches were performed on 30 June 2018 (search from start of databases) and an updated search on 22 March 2024 to identify studies from 2018 up to the present.

3.3. Results

A total of 984 potential articles were identified by the searches. After review against the inclusion and exclusion criteria, 40 studies were included in the final review (Figure 3.1).

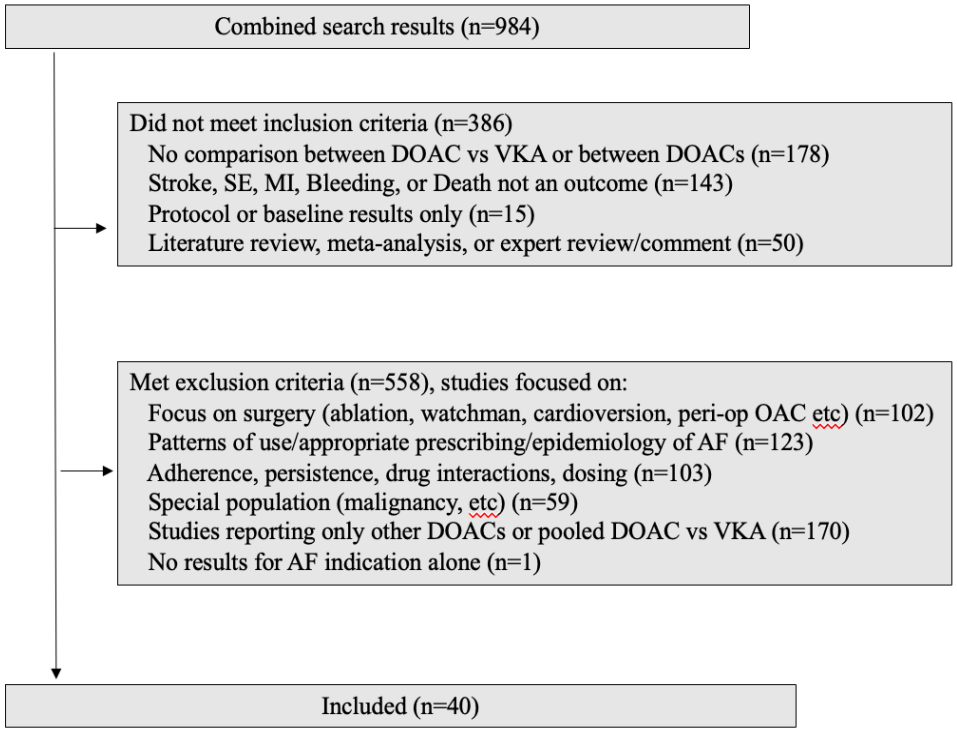


Figure 3.1 Flow diagram of search strategy and results

AF=atrial fibrillation; DOAC=direct-acting oral anticoagulant; MI=myocardial infarction; OAC=oral anticoagulant; peri-op=peri-operative; SE=systemic embolism; VKA=vitamin K antagonist.

3.3.1. Description of studies included in the final review

A total of 40 studies met the eligibility criteria and were included in the final review, 30 of which were identified at the time of the initial literature review (performed on 30 Jun 2018) and the remaining 10 identified during the course of completion of the thesis (updated search to identify studies from 01 July 2018 to present performed on 22 March 2024).

A list of included studies is given in Table 3.1. The studies used data ranging from 01 Oct 2010 (most studies used data starting from 01 Jan 2013) to Dec 31, 2019. Nineteen studies used data from the US (including US Department of Defence Data, Medicare, Medicaid, Marketscan, OptumInsight, PharMetrics, Optum Clinformatics and HealthCore US medical pharmacy claims), a further 16 studies used data from Western European countries (5 German [3 claims database, 2 outpatients IMS Disease Analyzer], 5 Danish nationwide registries, 1 UK [QResearch and CPRD], 2 Swedish EHRs, 2 Italian administrative database [1 from 10 local health units and 1 using linked data from the Emilia-Romagna region], 1 Spanish EHRs, and 1 French National Health System Claims Data [SNDS]), with the remaining 5 studies from East Asian populations (3 Japanese EHRs, 1 Korean and 1 Taiwanese National Health Insurance Research Database). The sample size of the apixaban group ranged from 723 to 110 259 with most studies (21) analysing between 1000 and 10 000 patients exposed to apixaban. Sixteen studies had more than 10 000 patients in the apixaban cohort of which two studies had more than 100 000 patients in the apixaban cohort, and only two studies had fewer than 1000 patients in the apixaban cohort. The two largest studies with more than 100 000 patients in the apixaban arm of the cohort pooled data from multiple US Claims datasets.

Most studies used propensity score matching (PSM) to control for confounding with inverse probability of treatment weighting (IPTW) and hazard ratios adjusted for baseline covariates also commonly used.

3.3.2. Summary of results of apixaban effectiveness and safety in the non-interventional studies

Table 3.1 summarises the data sources, study design, and key results, with studies grouped by geographical region.

Table 3.1 Summary of methods and results of the non-interventional studies included in the scoping review of effectiveness and safety of apixaban vs vitamin K antagonists

First author and year of publication	Data source Population	Methods used	Cohort size Apixaban/VKA Follow-up Key results
Region: North America (US)			
Yao X 2016 (82)	US **Insurance + Medicare Advantage** Patients with NVAF using OAC Oct 2010 to Jun 2015. Allowed prior warfarin exposure Excluded: valvular heart disease, ESRD, kidney transplant, or dialysis, hip or knee replacement 6 weeks prior, DVT or PE	PSM, sensitivity: TTR, ITT, censor at 6m, exclude ablation/cardioversion Censored at treatment switch or discontinuation No trial emulation	15 390/15 390 Max follow-up 4y 9m Stroke/SE 0.67 (0.46, 0.98) Major bleed 0.45 (0.34, 0.59)
Lip GYH 2016 (83)	US Truven **MarketScan + Medicare** supplemental claims New users with AF No information on excluded groups	PSM, no information on sensitivity No information on censoring No trial emulation	7438/ 7438, Max follow-up 3y Major bleed 0.53 (0.39, 0.71)
Lip GYH 2016 (84)	US Truven **MarketScan + Medicare** supplemental claims New users with AF Excluded: valvular heart disease, thyrotoxicosis, pericarditis, mitral stenosis, VTE, heart surgery and endocarditis	aHRs, sensitivity +30 days to censoring Censored at treatment switch or discontinuation No trial emulation	2402/ 12 713, Max follow-up 1y Major bleed 0.52 (0.30, 0.89)
Coleman CI 2016 (85)	US **MarketScan** claims New users with NVAF using OAC Jan 2012 to Oct 2014 Excluded: prior stroke, systemic embolism or ICH Funded by Bayer	PSM, no information on sensitivity No information on censoring No trial emulation	4083/ 4083 Max follow-up 2y 10m Stroke 0.63 (0.35, 1.12)
Li XS 2017 (86)	**4 US claims databases** (**MarketScan, PharMetrics, Optum, Medicare Advantage**) New users with NVAF using OAC Jan 2013 to Sep 2015 Excluded: valvular heart disease, VTE, transient AF (pericarditis, hyperthyroidism, thyrotoxicity), pregnancy	PSM, sensitivity: patients not censored 1 year post-index date Censored at treatment switch or discontinuation No trial emulation	38 470/ 38 470 Max follow-up 1y Stroke/SE 0.67 (0.59, 0.76) Major bleed 0.60 (0.54, 0.65)
Coleman CI 2017 (87)	US Truven **MarketScan (commercial + Medicare)** New users with NVAF using OAC Jan 2012 to June 2015 Excluded: Pts without previous ischemic stroke/TIA. Pts with transient cause NVAF, VTE, hip or knee arthroplasty, cancer, pregnancy, >1 OAC on index date or during follow-up	PSM, sensitivity shrunken cohort analysis (consistent results) Censored at treatment switch or discontinuation No trial emulation	1257/ 1257 Max follow-up 3y 6m, mean 0.5y Ischemic stroke/ICH 0.70 (0.33, 1.48) Major bleed 0.79 (0.38, 1.64)
Lin J 2017 (88)	**IMS Pharmetrics Plus**, US medical pharmacy claims No information on new/prevalent user inclusion. No information on excluded groups	PSM, no information on sensitivity No information on censoring No trial emulation	ND, Max follow-up 2y 9m warfarin vs apixaban Major bleed 2.05, (1.4, 3.0)
Adeboyeje G 2017 (89)	US **HealthCore** medical pharmacy claims New users with NVAF with OAC Nov 2010 to Feb 2015. Excluded: valvular heart disease, hyperthyroidism, DVT, PE, kidney transplant, dialysis, or hyperthrombotic conditions. Any OAC Rx in 6m prior to index date.	IPTW, no information on sensitivity Censored at treatment switch or discontinuation No trial emulation	3689/ 23 431 ND Major bleed 0.52 (0.41, 0.60)

First author and year of publication	Data source Population	Methods used	Cohort size Apixaban/VKA Follow-up Key results
Deitelzweig S 2017 (90)	US **Medicare Advantage** New users with NVAF using OAC Jan 2013 to Sep 2015 Excluded: age < 65 yrs, valvular heart disease, VTE, transient AF, cardiac surgery, hyperthyroidism, thyrotoxicity	PSM, no sensitivity Censored at treatment switch or discontinuation No trial emulation	14 214/ 14 214 Max follow-up 2y 9m Stroke/SE 0.65 (0.51, 0.83) Major bleed 0.53 (0.45, 0.63)
Hernandez I 2017 (91)	US 5% random sample of **Medicare** beneficiaries New users with AF Excluded: beneficiaries without continuous Part D enrolment	aHRs, no sensitivity Censored at treatment switch or discontinuation No trial emulation	2358/ 12 353 Max follow-up 1y Ischemic stroke/SE/death 0.86 (0.76, 0.98) Any bleed 0.79 (0.70, 0.90)
Lopes RD 2017 (92)	US **Medicare** New users with NVAF using OAC Jan 2013 to Dec 2014 Excluded patients without coronary/peripheral artery disease, age < 65 yrs	PSM, no information on sensitivity analyses No information on censoring at treatment switch or stop No trial emulation	9410/ 9410 Max follow-up 2y, mean 6m Stroke/SE 0.44 (0.30, 0.64) Major bleed 0.58 (0.49, 0.69)
Li X 2018 (93)	**4 US claims databases** (MarketScan, PharMetrics, Optum, Humana) New users with NVAF Excluded: pregnancy, valvular heart disease, VTE, transient AF (pericarditis, hyperthyroidism, thyrotoxicity)	PSM; Sensitivity using extended follow-up not restricted to 1 year (results consistent). Censored at treatment switch or discontinuation. No trial emulation	31 827/31 827 [5mg], 6600/6600 [2.5mg] Max follow-up 1y Stroke/SE 0.70 (0.60, 0.81) [5mg] Stroke/SE 0.63 (0.49, 0.81) [2.5mg] Major bleed 0.59 (0.53, 0.66) [5mg] Major bleed 0.59 (0.49, 0.71) [2.5mg]
Amin A 2018 (94)	US **OptumInsight** Research Database New users with NVAF using OAC Jan 2013 to Sep 2015 No information on excluded patients	PSM, no information on sensitivity analyses No information on censoring at treatment switch or stop No trial emulation	8328/ 8328 Max follow-up 2y 9m Warfarin vs apixaban: Stroke/SE 1.60 (1.23, 2.07) Major bleed 1.95 (1.60, 2.39) All-cause death 1.30 (1.21, 1.40)
Lopes RD 2018 (95)	US **Medicare** New users with NVAF using OAC Jan 2013 to Sep 2105 Excluded patients without coronary/peripheral artery disease, age < 65 yrs	PSM, no information on sensitivity analyses No information on censoring at treatment switch or stop No trial emulation	15 527/ 15 527; Max follow-up 2y 9m, mean 6m Stroke/SE 0.48 (0.37, 0.62) Major bleed 0.66 (0.58, 0.75)
Amin A 2018 (96)	US **Medicare + Medicaid** New users with AF with OAC Jan 2013 to Dec 2014 Excluded: age < 65 years, rheumatic mitral valvular heart disease, mitral valve stenosis, heart valve replacement or surgery; transient AF (pericarditis, hyperthyroidism, and thyrotoxicity), venous thromboembolism, pregnancy during the study period; or > 1 OAC prescription on index date.	PSM, sensitivity: i) included apixaban dose as subgroup and outcomes compared/interaction evaluated; ii) Patients censored at 6 months to create more balanced length of follow-up between groups; iii). Only patients with ≥ 30 days of follow-up evaluated. Censored at treatment switch or discontinuation No trial emulation	20 803/ 20 803 Max follow-up 2y, median 120 days warfarin: vs apixaban Stroke/SE 2.51 (1.92, 3.29) Major bleed 1.96 (1.71, 2.23)
Lip GYH 2018 (97)	US **Medicare, Medicaid, + 4 commercial claims databases**. New users with NVAF using OAC Jan 2013 to Sep 2015. Excluded: any OAC 12 months before index date, valvular heart disease, VTE, transient AF (pericarditis,	PSM within databases then pooled into 1 cohort, Sensitivity: i) Restricting follow-up to 1 year; ii) Multivariate Cox PH models on all patients meeting eligibility criteria (without PSM); iii) All-cause death for	100 977/ 100 977 Max follow-up 2y 9m, median 126 days apixaban, 158 days warfarin Stroke/SE 0.64 (0.58, 0.70)

First author and year of publication	Data source Population	Methods used	Cohort size Apixaban/VKA Follow-up Key results
	hyperthyroidism, thyrotoxicity), heart valve replacement/ transplant during baseline period, pregnancy, hip or knee replacement surgery within 6 weeks before index date.	Medicare patients only (other databases do not include complete death information) Censored at treatment switch or discontinuation. No trial emulation	Major bleed 0.60 (0.56, 0.63)
Gupta K 2019(98)	**US Department of Defence** NVAF patients with OAC Rx Jan 2013- Sep 2015. Excluded: valvular heart disease, heart valve replacement, dialysis, kidney transplant, end-stage chronic kidney disease, VTE, reversible AF, OAC during 12-months prior to index, hip or knee replacement within 6 weeks prior to index date, > 1 OAC on index date, or pregnancy	PSM. Sensitivity: i) cohorts stratified by dosage of DOACs (standard and reduced) on index date to assess if outcomes altered by DOACs dosage; ii) exclude patients with catheter ablation within 2 months prior to index, and exclude those with cardioversion 1 month before or after index; iii) max follow-up 6-months; iv) ITT Censored at treatment switch or discontinuation. No trial emulation	7607 / 7607 Max follow-up 2y 9m, median 161 days apixaban 153 days warfarin Stroke/SE 0.55 (0.39, 0.77) Major bleed 0.65 (0.53, 0.80)
Amin A 2020(99)	US **Medicare + Medicaid** Patients with AF (aged ≥ 65 years) initiating OAC Jan 2013 to Dec 2014. Excluded: rheumatic mitral valvular heart disease, mitral valve stenosis, heart valve replacement or surgery; transient AF, VTE, OAC during 12-month baseline period; pregnancy; > 1 OAC on index date.	PSM. Sensitivity: i) DOAC standard-dose and low-dose cohorts to look for interaction; ii) censored at 6 months; iii) restrict to patients with ≥ 30 days follow-up. Censored at treatment switch or discontinuation. No trial emulation	38 740 / 38 740 Max follow-up 2y, median 113 days warfarin, 97 days apixaban Warfarin vs apixaban Stroke/SE 2.18 (1.80, 2.64) Major bleed 1.76 (1.59, 1.95)
Franklin J et al 2021 [protocol] (100) Wang SV et al 2023 [results] (20)	**Pooled US Claims** Data RCT-DUPLICATE emulation of ARISTOTLE Marketscan: Jan 2013-Dec 2018, Optum: Jan 2013-Dec 2019. Medicare: Jan 2013-Dec 2017. ARISTOTLE eligibility criteria applied. No warfarin or apixaban prescription in the 180 days prior to index date.	PSM As-treated as primary analysis and ITT secondary analysis. Censored at treatment switch or discontinuation in primary as-treated analysis. Applied trial eligibility criteria but did not match to the trial on baseline characteristics.	110 259 / 110 259 Optum: 15 273 pairs Truven: 14 849 pairs Medicare: 80 137 pairs Max follow-up 1 year in ITT analysis. Median follow-up 98 days Stroke/SE on-trt 0.68 (0.61,0.76) Stroke/SE ITT sensitivity: 0.73 (0.67, 0.79)
Region: Europe			
Larsen TB 2016 (101)	**3 Danish nationwide databases** New users with AF Excluded: reduced dose apixaban, valvular AF or VTE (PE or DVT)	IPTW, sensitivity censoring at treatment switch Not censored at treatment switch or stop for primary - ITT used No trial emulation	6349/ 35 436 Max follow-up 4y 2m Ischemic stroke/SE 1.08 (0.91, 1.27) All-cause death 0.65 (0.56, 0.75) Major bleed 0.61 (0.49, 0.75)
Lamberts M 2017 (102)	**Danish administrative registries** New users with AF Excluded: recent (<6 months) VTE or PE or recent (<5 weeks) hip or knee prosthetic surgery	aHRs, no information on sensitivity Censored if >1 OAC Rx or no OAC Rx No trial emulation	7963/ 24 230 Max follow-up 4y 4m Major bleed 0.82 (0.70-0.95)
Nielsen PB 2017 (103)	**3 Danish nationwide databases** New users with AF	IPTW, sensitivity aHRs, restricted study period, supplementary subgroups	4400/ 38 893 Max follow-up 5y 7m

First author and year of publication	Data source Population	Methods used	Cohort size Apixaban/VKA Follow-up Key results
	Excluded: standard dose apixaban, valvular AF, PE, DVT, or recent hip/knee surgery	Not censored at treatment switch or stop for primary - ITT used No trial emulation	Ischemic stroke/SE 1.19 (0.95, 1.49) Major bleed 1.04 (0.76, 1.43)
Staerk L 2017 (104)	**Danish nationwide registries** New users with AF Excluded: age<30 or>100 yrs, valvular AF, hip or knee arthroplasties in prior 5 weeks, PE or DVT in 6 months prior	aHRs, sensitivity: with additional confounder, limited study period, ITT, and supplementary subgroup analysis Censored at treatment switch or discontinuation No trial emulation	6899/ 18 094 Max follow-up 2y Stroke/TE (excludes ICH): 1.07 (0.87, 1.31)
Coleman CI 2017 (105)	**German outpatients IMS Disease Analyzer** New users with AF Excluded: valvular AF, prior stroke, SE, or ICH, > 1 OAC on index date or OAC switch during follow-up	PSM, no sensitivity Not censored at treatment switch or stop, dealt with via exclusion criteria. No trial emulation	723/ 723 Max follow-up 1y 8m Ischemic stroke/TIA/MI/ICH 0.77 (0.43, 1.40) Major bleed 0.56 (0.34, 0.93) MI 0.28 (0.08, 0.99)
Hohnloser SH 2017 (106)	**German claims database.** New users with AF No information on excluded groups	aHRs, sensitivity: PSM No information on censoring No trial emulation	3633/ 16 179 Max follow-up 2y 3m Major bleed 0.68 (0.51-0.90)
Lip GYH 2017 (107)	**Danish nationwide registries** New users with AF Excluded: reduced dose apixaban, prior stroke, SE or TIA or age >75yr, >1 non-gender stroke risk factor, valvular AF, VTE (PE or DVT)	IPTW, sensitivity: on-treatment analysis censoring at treatment switch + falsification analysis Not censored at treatment switch or stop for primary - ITT used, censored for on-treatment sensitivity No trial emulation	1470/ 7674 Max follow-up 2y 6m Ischemic stroke/SE 1.01 (0.51, 2.03) Any bleed 0.35 (0.17, 0.72) All-cause death 0.47 (0.29, 0.76)
Coleman CI 2018 (108)	**German outpatients IMS Disease Analyzer** New users with AF Excluded: valvular AF, prior event in composite endpoint, >1 OAC on index date or OAC switch during follow-up, min 1 year follow-up	PSM, no sensitivity Not censored at treatment switch or stop, dealt with via exclusion criteria. No trial emulation	835/ 835 Max follow-up 2y 3m Ischemic stroke/TIA/MI/ICH 0.87 (0.47, 1.60) Ischemic stroke 1.51 (0.54, 4.24) MI 0.33 (0.11, 1.03)
Hohnloser SH 2018 (109)	**German claims database** No information on whether new or prevalent users. No information on exclusions.	aHRs, sensitivity: PSM No information on censoring No trial emulation	10 117/ 23 823 No information on duration of follow-up Apixaban vs Phenprocoumon Stroke/SE 0.77 (0.66, 0.90) Major bleed 0.58 (0.49, 0.69)
Sjalander S 2018 (110)	**Swedish EHRs, identified through Auricula** New users with AF, no information on exclusions TTR 0.714 in warfarin users	Matching on Mahalanobis No information on censoring at treatment switch or stop. No trial emulation	12 311/ 12 311 Max follow-up 3y Stroke/SE 0.92 (0.70, 1.20) Major bleed 0.63 (0.52, 0.75) All-cause death 0.83 (0.72, 0.96) MI 0.68 (0.49, 0.95)
Vinogradova	**UK QResearch + CPRD Gold linked to HES and ONS**	aHRs, sensitivity: IPTW	10 601/ 70 585.

First author and year of publication	Data source Population	Methods used	Cohort size Apixaban/VKA Follow-up Key results
Y 2018 (111)	2011-2016 New users with AF with OAC in study period [QResearch: Jan 2011 to Oct 2016 CPRD Gold: Jan 2011 to Mar 2016] Inclusion: aged 21 to 99 at study entry (entry date first Rx of any anticoagulant). Excluded: any OAC Rx in 12 months before entry; <12 months records before entry; no valid Townsend score.	Censored at treatment switch or discontinuation No trial emulation. Prospective open cohort study, looked at DOACs vs warfarin and risks of bleeding, ischemic stroke, VTE, and all-cause mortality.	Max follow-up 5y 9m Ischemic stroke 1.13 (0.89,1.44) ICH 0.40 (0.25,0.64) Major bleed 0.66 (0.54, 0.79) All-cause death: 1.13 (1.01, 1.25) [all doses] 0.98 (0.83, 1.15) [5mg apixaban dose] 1.27 (1.12, 1.45) [2.5mg apixaban dose]
Ramagopalan SV 2018 (112)	**Italy administrative database 10 local health units** New users with AF Excluded: valvular AF, VTE (DVT/PE), knee or hip replacement in prior 6 weeks, more than one OAC	PS quintiles, no information on sensitivity Censored at treatment switch or discontinuation No trial emulation	1521/ 8393 Max follow-up 3y 5m Major bleed 0.44 (0.20- 0.97)
Marietta M 2019 (113)	**Italian administrative healthcare datasets** linked to data gathered in study centres from 7 anticoagulation clinics (Emilia-Romagna Region). Adults with NVAF enrolment Oct 2014 to Jan 2017. Included both patients naïve to any anticoagulant treatment and those already on VKA or DOAC (3 categories: 'VKA experienced' if >90 days VKA exposure in the 180 days before index date; 'DOAC experienced' if >90 days DOACS in the 180 days before index date, 'Naïve' in all the remaining cases. Index date: first prescription of OAC after the enrolment.	PS-adjusted HRs calculated for pooled DOAC comparisons only. Competing risk approach used – Fine & Gray proportional hazards models to assess relationship between treatment and outcomes Censored at treatment switch or discontinuation. No trial emulation	1 955 / 954 Stroke/SE crude HR 0.90 (0.45, 1.79) Major bleed crude HR 1.50 (0.92, 2.44) All-cause death crude HR 1.28 (0.75, 2.21) Mean TTR in VKA 74.0
Ramagopalan SV 2019 (114)	**Spanish EHR** records from primary and hospital centres. NVAF patients, with Rx for apixaban or VKA (acenocoumarol or warfarin) Jan 2015 to Dec 2017. Patients prescribed apixaban who had switched from VKA and vice versa eligible for inclusion. Excluded: transferred to other centres, displaced or out of area; permanently institutionalized; history of transient AF (thyrotoxicosis, pericarditis), heart surgery, VTE, hip or knee surgery in previous 6 weeks, valvular heart disease and/or pregnancy; subjects with valvular AF (with prosthetic valves); and end-stage renal disease, dialysis or kidney transplant.	PSM, no info on sensitivity Censored at treatment switch or discontinuation. Patients who started VKA and switched to DOAC other than apixaban during follow-up excluded. No trial emulation	2160 / 2160 (acenocoumarol as the VKA) 28.2% of apixaban VKA-experienced, all of VKA arm VKA-experienced. Max follow-up 4y Stroke/SE 0.54 (0.38, 0.78) Major bleed 0.51 (0.37, 0.72)
Van Ganse E 2020 (115)	**French National Health System claims data** (SNDS, note SNDS did not have clinical history, clinical or paraclinical examination (tobacco smoking, blood pressure level, BMI, etc), lab results). Adult patients with NVAF using OAC Jan 2014 to Dec 2016. AF diagnosed in the 2yrs prior to index date (first Rx in study period).	PSM 1:n; Sensitivity: i) matching on high-dimensional PS; ii) adjustment on PS; and iii) adjustment on known confounders. Cumulative incidence rates account for competing risk of death using Fine and Gray models. Censored at treatment switch or discontinuation. No trial emulation	68 208 / 107 558 Max follow-up 3y, median 218 days for VKAs, 213 days for apixaban. Stroke/SE 0.60 (0.56–0.65) Major bleed 0.43 (0.40–0.46) All-cause death 0.44 (0.42, 0.45)

First author and year of publication	Data source Population	Methods used	Cohort size Apixaban/VKA Follow-up Key results
	Excluded: multiple OAC Rx on index date, multiple doses or prescribers at index date, patients possibly treated for indications other than stroke prevention in NVAF, any use of the same OAC in the 2 yrs prior to index date .	Additional sensitivity using modified definitions of outcomes: addition of transfusion for bleeding; and for effectiveness outcome, exclusion of haemorrhagic stroke.	
Warkentin L 2023(116)	**German health insurance data** NVAF patients with a first prescription for OAC 2015–2018 (no OAC Rx in prior 12 months), Excluded: warfarin as VKA (only phenprocoumon included), reduced-dose DOAC, patients with >1 OAC or different doses at index, patients with < 12 months follow-up (unless died)	aHRs as primary, PSM as sensitivity Stroke/SE outcome excluded ICH Censored at treatment switch or dose change or discontinuation. No trial emulation Discussion notes PREFER IN AF study showed patients with phenprocoumon higher TTR (79%)(117)	23 343 / 20 179 (14 939 / 14 939 for PSM analysis) Thromboembolic event 1.08 (0.94, 1.25) Death: 0.95 (0.87, 1.05) Major bleed: 0.54 (0.46, 0.63) All-cause death 0.95 (0.87, 1.05)
Region: Asia			
Kohsaka S 2017 (118)	**Japanese 275 hospital EHRs** New users with AF No information on excluded groups	PSM, no information on sensitivity No information on censoring No trial emulation	5977/ 5977 Max follow-up 1y Major bleed 0.59 (0.42-0.82)
Cha MJ 2017 (119)	**Korean National Health Insurance Service database** New and prevalent users with AF (subgroup analysis) Excluded: stroke, TIA, or ICH within 10-years prior to study start	PSM, no sensitivity, supplementary subgroup analyses Not censored for discontinuation – ITT used but switchers excluded No trial emulation	2189/ 4378 Max follow-up 2y Ischemic stroke/ICH 0.51 (0.28- 0.82) All-cause death 0.32 (0.18, 0.53)
Chan YH 2018 (120)	**Taiwan National Health Insurance Research Database** New users with AF, Excluded: valvular AF, VTE (PE or DVT) or joint replacement 6m prior, ESRD, age <30 yrs	IPTW, no sensitivity analyses Did not censor at treatment switch or discontinuation No trial emulation	5 843/ 19 375; 4y 7m max follow-up Ischemic stroke/SE 0.55 (0.43, 0.69) Major bleed 0.41 (0.31, 0.53) All-cause death 0.58 (0.51, 0.66)
Kohsaka S 2018 (121)	**Japan EHRs claims data from 314 acute-care hospitals** AF patients newly initiated (no prescription during the 180-day blanking period) Excluded: valvular AF, postoperative AF, hyperthyroidism or thyrotoxicosis, ESRD or pregnancy	PSM, no sensitivity analyses Censored at treatment switch or discontinuation No trial emulation	11 972/ 11 972 Max follow-up 6y 4m Stroke/SE 0.64 (0.48, 0.85) Major bleed 0.66 (0.51, 0.85)
Kohsaka S 2020 (122)	**Japanese health claims from 372 acute care hospitals** (Medical Data Vision Co Ltd). Adult patients with NVAF, OAC Mar 2011 to July 2018 and no OAC during baseline (year preceding index date). Excluded: valvular AF, postoperative AF, AF associated with mechanical valve malfunction or complication of heart valve prosthesis or rheumatic AF, hyperthyroidism or thyrotoxicosis, procedures involving prosthetic heart valves during baseline period, dialysis, pregnancy.	IPTW. Sensitivity: i) restricting follow-up to 1 year, ii). 1:1 PSM used. Censored at treatment switch or discontinuation. No trial emulation	22 336 / 15 902 Max follow-up 2y Stroke/SE 0.65 (0.558, 0.766) Major bleed 0.72 (0.614, 0.843)

AF=atrial fibrillation; aHR=adjusted hazard ratios; BMI=body mass index; DOAC=direct-acting oral anticoagulant; DVT=deep vein thrombosis; EHR=electronic healthcare record; ESRD= End stage renal disease; ICH=intracranial haemorrhage; IPTW=inverse probability of treatment weighting; ITT=intent-to-treat; max=maximum; MI= myocardial infarction; OAC=oral anticoagulant; m=month; ND=no data; NVAF=non-valvular atrial fibrillation; PE=pulmonary embolism; PSM=propensity score matching; Rx=prescription; SE = systemic embolism; TIA=transient ischemic attack; VKA=vitamin K antagonist; VTE=venous thromboembolism; y=year. Valvular heart disease=rheumatic mitral stenosis, mechanical or bioprosthetic heart valve, or mitral valve repair.

3.3.2.1. Methods comparisons

All of the studies used suitable methods such as PSM, IPTW, or adjustment of the HRs to control for confounding for their primary analysis (the only study that presented crude HRs only for the apixaban vs VKA comparison(113) did use suitable methods for their primary analysis involving pooled DOACs).

Twenty-two of the studies reported censoring patients that switched or stopped treatment to estimate 'on-treatment' estimates; using this approach alone could result in biased estimates if there were any differential switching conditional on intercurrent events. For example, should a patient develop a contra-indication to DOACs during follow-up (such as a diagnosis of mitral stenosis, placement of a mechanical heart valve, or severe renal impairment) then such patients would be more likely to switch treatment from a DOAC to a VKA and be censored at this timepoint compared with users of VKA with the same diagnosis that could continue on their index treatment. Many of the contra-indications to DOAC treatment are associated with an increased risk of ischemic and/or bleeding events meaning censoring at treatment switch would be informative censoring potentially leading to biased estimates making warfarin look worse. In addition, one study (116) censored patients if they had a change in the dose of their index DOAC which could also lead to bias given that the indications for DOAC dose reduction (such as old age and reduced kidney function) are related to outcomes of interest.

Acting in the opposite direction there is a risk of patients doing badly on warfarin (for example those with minor bleeding or low TTR) being more likely to switch to a DOAC during follow-up which could lead to informative censoring of users of warfarin leading to an overestimate of the relative benefit of warfarin. Such switching of warfarin users may also

impact any ‘intent-to-treat’ analyses if the warfarin users that switched experienced lower event rates on the DOAC than they would have experienced had they stayed on warfarin.

The relative impact of these two directions of bias resulting from treatment switching depends on the relative occurrence of these events during follow-up and it is difficult to ascertain the potential impact without summaries of the timing and reason for switching. Seven studies reported performing sensitivity analyses using a shorter follow-up time limit (82)(91)(92)(96, 97, 104)(115); such an approach could help ascertain the impact of treatment switching on treatment effect estimates depending on the distribution of time to treatment switch.

Some studies (101) (114) attempted to remove the impact of treatment switching by excluding patients that switched treatment during follow-up; this could lead to selection bias as treatment switching is likely related to outcomes.

3.3.2.2. Primary efficacy endpoint comparisons

Twenty non-interventional studies reported the primary efficacy endpoint of stroke/SE, of which 6 showed results consistent with the reduced risk with apixaban compared with warfarin demonstrated in ARISTOTLE (using the non-interventional study 95% confidence intervals (CI) intersecting the confidence interval from ARISTOTLE HR 0.79, 95% CI [0.66–0.95] as the definition of ‘consistent’). Five studies reported larger treatment effects not consistent with the 95% CI of ARISTOTLE (with point estimates for the hazard ratios ranging from 0.40 to 0.60, though notably 4 of these were in patients aged 65 or older using US Medicare (92, 95), or Medicare and Medicaid combined (96, 99), and the other study used French EHRs (115). Other than 2 studies (HRs 0.92 in Swedish EHRs (110) and 0.90 in Italian EHRs (113)) all hazard ratios were lower (range 0.40 to 0.77) than the ARISTOTLE estimate of 0.79 indicating the non-interventional studies were detecting a larger protective

effect than the trial. Five studies reported a different endpoint of ischemic stroke or SE all finding no significant difference between apixaban and warfarin (HRs ranged from 1.01(107) to 1.19(103)) comparable to the ARISTOTLE endpoint of ischemic or uncertain type of stroke which showed no difference (0.92 [0.74–1.13]). These results reflect that the primary reduction in the stroke/SE endpoint comes from the superior safety of apixaban (lower haemorrhagic stroke risk) rather than superior efficacy. Other possible explanations for the higher relative risk for ischemic stroke on apixaban in the non-interventional studies include underdosing where patients suitable for standard dose apixaban are prescribed the reduced dose (for example Steinberg et al found 9.4% of NOAC patients were underdosed in an AF registry study(123), and Harrington et al 2022 found 61% of patients prescribed reduced-dose apixaban did not meet the criteria for dose reduction, and these patients were at higher risk of ischemic stroke (124)), and the possible lower drug adherence in real life compared with trial conditions (a study in US claims data by Brown et al found adherence to apixaban of 82% as proportion of days covered by prescriptions (125))

3.3.2.3. Primary safety endpoint comparisons

Thirty-two studies reported results for the primary safety endpoint of major bleeding with most broadly comparable with the ARISTOTLE result (0.69 95% CI [0.60–0.80]); all but 3 studies reported hazard ratios lower than ARISTOTLE (HRs ranged from 0.41 to 0.79) and only 4 studies reported results (0.41 [0.31–0.53], 0.45 [0.34–0.59], 0.51 [0.44, 0.58], 0.43 [0.40, 0.46]) not consistent with ARISTOTLE (all suggesting a greater safety margin for apixaban than detected in ARISTOTLE). All other bleeding endpoints for which there were non-interventional study estimates (intracranial, other location, gastrointestinal, and any bleeding) also showed similar lower risk for apixaban vs. warfarin as expected from ARISTOTLE.

3.3.2.4. Secondary effectiveness endpoints comparisons

Results for other key outcomes including MI and all-cause death showed a lower risk for patients treated with apixaban versus VKAs consistent with the trial but with the effect size sometimes exceeding that shown in the trial. Several studies reported much larger reductions in the risk of all cause death on apixaban vs warfarin (for example HRs of 0.32, 0.44, 0.47, and 0.58 vs ARISTOTLE HR 0.89). Two studies suggested a benefit for apixaban vs warfarin in excess of that observed in ARISTOTLE for the risk of MI (HRs of 0.28 and 0.33 vs ARISTOTLE HR 0.88). Potential reasons for these studies showing larger reductions in the risk of death and MI on apixaban vs warfarin when compared with the ARISTOTLE results are given in the next section.

The only outcome other than ischemic stroke or SE to show no benefit for apixaban over warfarin across multiple studies was ischemic or uncertain type of stroke. In ARISTOTLE there was no difference for this outcome (0.92 [0.74–1.13]); likewise, in the non-interventional studies no difference was found between treatments with all confidence intervals crossing 1 other than 2 studies that showed lower risk for apixaban. Five studies reported a HR showing a higher risk of ischemic stroke with apixaban (estimates ranged from 1.11[10] to 2.04[14]) though in all cases this difference was not significant.

3.3.3. Summary of methods and key characteristics of the non-interventional studies

Table 3.1 summarises the methods used and key characteristics of the studies. Most (n=23) studies used PSM to control for confounding, inverse probability of treatment by propensity score weighting was used in 6 studies and calculation of hazard ratios adjusted for baseline covariates in 8 studies. One study used Mahalanobis distance matching and one study used random sampling of VKA cases within propensity score quintile strata to obtain the same

number of VKA as apixaban users within each stratum, with adjusted hazard ratios. One study (Marietta 2019(113)) presented crude HRs only for the individual DOAC vs VKA comparisons though did use PS methods for the study's primary comparison which involved pooled DOACs.

One of the studies evaluated the potential for residual confounding by performing a 'negative control analysis' using outcomes expected not to be associated with treatment (such as pneumonia and hip fractures). Negative control outcomes are assumed to have i) no causal relationship with the treatments being compared and ii) to have the same confounding structure as the treatments and main study outcome of interest. Analysis of negative controls can be used in non-interventional studies to detect potential bias that may have occurred due to unmeasured confounding (since we expect to see no association between the treatments and the negative outcome if there is no confounding) (126); however it can be difficult to assess the validity of the assumption that the negative outcome selected shares the same confounding structure as the treatments and outcome of interest. In the studies which used negative outcome analysis the negative outcomes did not falsify (the a priori null hypothesis of neutral associations was generally rejected) indicating possible residual confounding after applying the inverse probability of treatment weighting.

The five studies that found lower point estimates for the relative risk of stroke or SE (0.44, 0.48, 0.40, 0.46, and 0.60) with CIs that did not overlap with ARISTOTLE (upper limits of 0.64, 0.62, 0.52, 0.56, and 0.65 vs ARISTOTLE 95% CI 0.66–0.95) were all in new users of OACs and mostly conducted in patients aged over 65 years using US data (4 of the studies, 2 in US Medicare and 2 in US Medicare and Medicaid combined) with 1 study using French EHRs. The greater treatment effect observed may have been due to this being a higher risk population and the warfarin arm consisting solely of new users with the heightened risk of

stroke and bleeding around the time of warfarin initiation, with a study by Azoulay et al 2014 in CPRD data showing a 71% increased risk of stroke in the first 30 days post-initiation of warfarin compared with non-use whereas a 50% decreased risk was observed >30 days post-initiation.

Four studies found a reduced risk of major bleeding (HRs 0.41(120), 0.45 (82), 0.51 (96), 0.43(115)) with CIs that did not overlap with ARISTOTLE (upper limits of 0.53, 0.59, 0.58, and 0.46 vs ARISTOTLE 0.69 [0.60-0.80]). The Chan study(120) was conducted in Asian people (Taiwan); previous studies have shown Asian people may be more sensitive to warfarin leading to a higher risk of intracranial haemorrhage (ICH) than people of other ethnicities (127) and subgroup analyses of pivotal trials in DOAC in Asian people showed that DOACs displayed better efficacy and safety in Asian people than in people of other ethnicities (128). The Chan study used inverse probability score weighting however it is also possible that residual confounding by unmeasured variables and selective prescribing could have contributed to the treatment effect observed. Patients were censored at first occurrence of any study outcome which may have biased the results if one treatment was more likely to cause one type of event as other events would not be detected. The Yao study (82) was conducted using US insurance and Medicare Advantage patients and used an on-treatment analysis approach censoring patients at discontinuation or switch of treatment. The Amin study (96) also used US data (Medicare and Medcaid) and was restricted to patients aged ≥ 65 years, a population which may be at higher risk of bleeding on warfarin. Finally the Van Ganse study (115) was a large study using French EHRs, the higher risk of bleeding on warfarin in this study compared to ARISTOTLE may be attributable to the older age of the cohort and to lower quality of INR control in France (Cotte et al 2014 in a study using European primary care databases found 52% of patients with NVAF on VKA in France had a TTR ≤ 70% compared with only 35% of patients in the UK (129)).

Both studies that reported a stronger protective association between apixaban and MI (HRs of 0.28 and 0.33) involved small sample sizes (723 and 835 PSM pairs respectively) and therefore had relatively wide confidence intervals which covered the possibility of a result consistent with ARISTOTLE (upper limit of 95% CI 0.99 and 1.03 respectively versus ARISTOTLE HR estimate of 0.88). In both studies patients who switched to another OAC during the 1-year follow-up period were excluded from the cohort prior to propensity score matching. This exclusion may have introduced selection bias into the studies, for example if the patients that were healthier during the follow-up were more likely to switch from warfarin to apixaban then this approach would be dropping the warfarin patients less likely to experience an outcome whereas those with contra-indications to DOACs (such as worsening kidney function) would be more likely to be selected into the warfarin arm of the cohort. Both these studies and another study[13] which reported a stronger protective association with all cause death (0.32 vs ARISTOTLE 0.89) excluded patients who had a history of thromboembolic events or ICH with study [24] also excluding patients with prior MI; in ARISTOTLE approximately 20% of patients had prior stroke or TIA.

For the outcome of death from any cause ARISTOTLE reported an 11% lower risk for apixaban vs warfarin [0.89 (0.80, 0.998)]. Eight of the non-interventional studies reported this outcome with 5 of these showing a protective effect of apixaban in excess of that predicted by ARISTOTLE ranging from a 35% reduction to a 68% reduction in risk of death (HRs of 0.32 (119), 0.47 (107), 0.58 (120), 0.65 (101), and 0.44 (115) with corresponding upper limits of the CI of 0.53, 0.76, 0.66, 0.75, and 0.45); these treatment estimates came from a range of different countries (Korea, Denmark, Taiwan, Denmark, and France respectively). It is difficult to ascertain to what extent these estimates represent real causal differences in benefit compared with the RCT due to differences in the relative risks and benefits in the populations included in these studies vs limitations from the methods used

such as residual confounding or bias contributing to the differences. A comparison of the causes of death between the treatment arms would be useful to aid understanding of the cause of these large differences in death rate – if most of the excess death in the VKA arms was caused by death due to stroke or bleeding complications this may suggest the increased benefit observed was a true effect and not caused by bias or confounding.

The most relevant study for the objective of this thesis was the emulation of ARISTOTLE using US claims data by the RCT-DUPLICATE initiative. Wang et al reported results from a large cohort of ARISTOTLE-eligible patients from several US Claims databases. The patients were propensity score matched enabling balance across a high number of covariates. Wang et al (16) found results consistent with ARISTOTLE though they found a slightly larger benefit of apixaban vs warfarin for stroke/SE based on the point estimates (on-treatment 0.68 (0.61, 0.76) and ITT sensitivity 0.73 (0.67, 0.79) in RCT-DUPLICATE vs 0.79 (0.66, 0.95) in ARISTOTLE ITT). A comparison of the RCT-DUPLICATE cohort against the ARISTOTLE participants reveals differences in the proportion of patients aged ≥ 75 years (59% vs 31% in ARISTOTLE), and the proportion of female patients (52% vs 35% in ARISTOTLE), ethnicity, and concomitant medications at baseline (Table 3.2). RCT-DUPLICATE excluded patients with warfarin or apixaban prescriptions in the 180 days prior to index date meaning continuing users of warfarin were not included in the cohort whereas ARISTOTLE had 57% of participants that were VKA-experienced (and had been randomised to continue on warfarin or switch to apixaban). VKA-experienced patients may be expected to have superior control of their warfarin therapy compared to new users given the time taken to find an optimal dose for a patient and familiarity with the lifestyle adjustments required on VKA therapy (such as consideration of food and alcohol interactions with VKA). The RCT-DUPLICATE ARISTOTLE emulation had a relatively short median follow-up time of 98 days.

Table 3.2 Comparison of baseline characteristics of ARISTOTLE RCT and RCT-DUPLICATE emulation of ARISTOTLE

Characteristic - % unless specified	ARISTOTLE Apixaban (N = 9 120)	ARISTOTLE Warfarin (N = 9 081)	RCT-DUPLICATE Apixaban (N = 110 259)	RCT-DUPLICATE Warfarin (N = 110 259)
Age – years , mean (SD)	69.1 (9.61)	69.0 (9.74)	76.1 (8.98)	76.2 (8.92)
Female sex	35.5%	35.0%	52.2%	52.3%
Qualifying risk factors				
Age ≥ 75 years	31.2%	31.1%	59.0%	59.4%
Diabetes	25.0%	24.9%	Approx. 25%	Approx. 25%
Hypertension req. treatment	87.3%	87.6%	Approx. 87.1%	Approx. 87.2%
Medications at index date				
Digoxin	32.0%	32.1%	10.9%	14.3%
Statin	45.0%	45.1%	55.5%	55.4%
Non-steroidal anti-inflammatory	8.2%	8.5%	11.5%	11.4%
Ethnicity				
White	82.6%	82.5%	92.7%	92.7%
Black	1.4%	1.1%	3.5%	3.5%
Asian	14.4%	14.7%	1.0%	1.0%
Other	1.6%	1.7%	2.8%	2.8%

Source: C B. Granger et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2011; 365:981-992, doi: 10.1056/NEJMoa1107039
SD = standard deviation

3.3.3.1. Impact of data source

The studies identified used a wide range of different non-interventional data sources, there are differences between commercial claims data and EHR data that have an impact on the interpretation of the treatment effect estimates. Electronic healthcare records tend to have a long follow-up of patients and depending on the country (eg UK and Sweden) may have data from birth on the patients included in the databases providing a rich source of information for the derivation of covariates. Healthcare systems like the UK's NHS lead to data sources that should be representative of the patient population in that country whereas commercial insurance-based data sources may only represent a younger subset of the population with higher socioeconomic status. Some data sources such as US Medicare contain data from only patients aged 65 and older or younger patients with certain conditions. Thus the different data sources can contain different subsets of the full patient population for the target indication.

The data source impacts the sample size with the nationalised health services such as in the UK resulting in large datasets whereas some countries have a more fragmented healthcare system leading to smaller datasets such as the Italian regional studies.

Whereas in EHRs the filling of a prescription may not be captured (such as in the UK data sources), the insurance claims data does provide this information which may increase the likelihood of accurate ascertainment of exposure. US insurance claims data sources can have relatively short duration of follow-up if patients are lost from the system when they change provider. Some data sources contain rich lifestyle information and demographic data, such as UK data which typically has ethnicity, smoking, and alcohol consumption recorded. Furthermore, differences in healthcare infrastructure between difference geographical regions can have a large impact, particularly in this therapeutic area in which the quality of warfarin therapy has been shown to vary widely between countries. Ethnicity can differ between the data sources adding a further source of variation to consider.

3.4. Conclusions

All of the non-interventional studies reported a lower risk of the primary efficacy (stroke/SE) and safety (major bleeding) outcomes for patients on apixaban compared with VKA with most studies reporting larger treatment effects than those observed in ARISTOTLE. Most of the non-interventional studies results were consistent for the primary efficacy and safety outcomes with only a few studies reporting hazard ratios and confidence intervals that did not overlap with the ARISTOTLE results. There were some key differences between the study designs and the ARISTOTLE trial design. Most notably the majority of studies used a new user design whereas in the trial 57% of subjects had prior exposure to a VKA; this may have impacted the results given the expected time to achieve good INR control following VKA initiation. Only one of the studies used UK data meaning that the proposed project will provide useful evidence on the effectiveness of apixaban compared with warfarin in UK

clinical practice and provide insight into the consistency of treatment estimates from different studies and methods using data from the same country.

Some studies excluded patients who switched treatment during the first year after initiation of apixaban or warfarin. Excluding such patients introduces a risk of biasing the results as it is possible that patients who switched treatment were those not tolerating the drug.

Whilst most studies used inclusion and exclusion criteria broadly similar to those used in ARISTOTLE only 1 study attempted to emulate ARISTOTLE by applying the trial eligibility criteria. None of the studies matched from the RCT population to the non-interventional cohort as planned by this project; the RCT-DUPLICATE emulation of ARISTOTLE used a cohort of patients that differed from the ARISTOTLE participants on age, sex, and many other characteristics. By matching to the RCT population and following an analysis approach as similar as possible to that used in the trial the results obtained should elucidate the difference between the UK real world effectiveness compared with the trial efficacy. RCT-DULICATE did not follow the successful benchmarking of their results to extend the study to patient groups underrepresented in or excluded from ARISTOTLE which is an analysis that can be explored in this project.

Most studies used propensity score matching to control for confounding. As part of this project propensity score matching and other methods will be explored.

Only 1 study tried to emulate the ARISTOTLE trial by applying the same eligibility criteria while many other studies used similar exclusion criteria. None of the studies matched to the trial participants. Many of the studies compared the results obtained with the ARISTOTLE results as part of the discussion section, however no studies other than the RCT-DUPLICATE emulation of ARISTOTLE assessed the validity of the results against a set of prespecified

criteria. RCT-DUPLICATE did not extend the analysis to underrepresented groups following benchmarking.

3.5. Limitations of review

This review was limited to studies including apixaban, the methods used in studies reporting results for pooled DOACs groups and other individual DOACs may also have relevance to the overall project though the methods are likely to be similar to the studies included in this review. References may have been missed, however, the search was performed on multiple databases, a grey literature search was performed, and the search updated to ensure no recent studies were missed.

3.6. Summary

This chapter summarised the results of the scoping literature review looking at non-interventional studies comparing apixaban against warfarin for the prevention of stroke in patients with AF. I found that whilst there have been many non-interventional studies looking at apixaban effectiveness only 1 applied the ARISTOTLE eligibility criteria, no studies matched to the baseline characteristics of the ARISTOTLE trial population, and only 2 studies included prevalent users of warfarin, and no studies extended the analysis after performing benchmarking against ARISTOTLE. The only study that emulated ARISTOTLE and benchmarked their results did not extend the analysis to under-represented or excluded patient groups. The wide range of treatment estimates observed showed the difficulty in drawing conclusions from non-interventional studies and suggested differences between countries as being important given observed differences in treatment risks and benefits by ethnicity and differences in the quality of VKA therapy between countries. The approach taken to account for treatment switching during follow-up is also likely to have contributed to the differences in treatment estimates seen given the risk of informative censoring for these

treatments in this indication. The existence of treatment switching associated with i) development of contraindications to DOAC therapy and ii) increased likelihood of switching from VKA to DOAC in cases where a patient has evidence of doing badly on warfarin means this bias could act in two possible directions. The next chapter will present the methods for the creation of an ARISTOTLE-analogous cohort of patients in CPRD Aurum.

Chapter 4 Methods for the emulation of ARISTOTLE in CPRD Aurum

This chapter will describe the methods employed to emulate the ARISTOTLE RCT using EHR data (CPRD Aurum linked to ONS and HES) including the protocol(1) followed by more detailed information on the methods:

- The protocol paper published for this study which summarises the planned methods at a high-level and serves as an introduction to this chapter.
- Determination of benchmarking criteria for the trial emulation.
- Derivation of code lists, treatment windows, algorithms used in baseline covariate classification, and application of the trial eligibility criteria to the cohort.
- Additional information on the methods used in selecting a subset of apixaban users in CPRD Aurum matching the ARISTOTLE trial apixaban arm participants
- Additional information on the method chosen for inclusion of prevalent users of warfarin.
- Analysis of outcome measures and sensitivity analyses.

A later chapter (Chapter 6) will cover any modifications to the methods used and results obtained in looking at special patient populations.

4.1. Paper 1: Real world effects of medications for stroke prevention in atrial fibrillation: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results

Emma Powell[1], Ian Douglas[1], Usha Gungabissoon[2], Liam Smeeth[1], Kevin Wing[1]

[1]Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

[2]GlaxoSmithKline, London, United Kingdom

4.2. Introduction to Paper 1

Summary

Chapter 3 reviewed existing non-interventional studies assessing effectiveness of apixaban compared to VKA therapy in terms of the populations studied, methods used, and results, and how these compared with the ARISTOTLE trial. It was noted only 1 study (RCT-DUPLICATE) tried to emulate the ARISTOTLE trial and assess the validity of the results against a set of prespecified criteria. The RCT-DUPLICATE emulation of ARISTOTLE did not match to the baseline ARISTOTLE patient characteristics, used US claims data, excluded prevalent users, and had a relatively short median follow-up time. In this chapter a protocol for the proposed study using CPRD Aurum and Gold data linked to HES and ONS to emulate the ARISTOTLE trial is presented. The protocol paper summarises the objectives of the study, data sources used, study design, validation criteria, and methods. The protocol paper was published in April 2021 in *BMJ Open*; in the paper the term 'replicate'/'replication' was used as this was the terminology being used early on in this PhD, throughout the rest of the thesis and in the results paper the term 'emulate/emulation' is used instead to align with the terms most commonly used by other researchers in this area. Where 'replicate/replication' is used in the protocol paper this can be considered to equate with the term 'emulate/emulation'.

Thesis objectives addressed

This chapter describes the analyses that were planned in addressing the following objectives of the overall thesis (Section 1.3):

1. Emulate the reference trial ARISTOTLE comparing apixaban to warfarin for prevention of stroke in atrial fibrillation in UK EHRs including application of the trial eligibility criteria, matching to the baseline characteristics of the participants in the reference trial, and assessing the validity of the results and methods by benchmarking.

2. To explore different methods in the emulation of the reference trial including different methods of matching and the inclusion of prevalent users.

Role of candidate

I drafted the paper providing a summary of the planned analyses, I wrote the ISAC protocol this paper was based on and a more detailed Statistical Analysis Plan submitted to Bristol-Myers Squibb in the application for the individual patient trial data. Kevin Wing (KW) provided an example protocol previously published in the BMJ Open based on the emulation of a COPD trial (130); his work helped inform the rationale for this PhD along with guiding the high-level plan of this trial emulation work and served as an example on what details to include in the protocol paper. Ian Douglas (ID), Usha Gungabissoon (UG) and KW provided guidance on the data sources, methods, and the strengths and limitations of the proposed study.

I performed the feasibility work involved in developing the protocol including review of the RCT protocol, development of appropriate codelists to apply the trial eligibility criteria to the EHR databases, and application of trial criteria to a sample of CPRD Gold data to estimate the cohort sample size. ID, UG, and KW advised on the implementation of the trial eligibility criteria. Liam Smeeth (LS) provided guidance as a clinician in the development of some of the codelists in particular in terms of which recorded symptoms and diagnoses a GP may consider clinically relevant in assessing a patient's risk of bleeding. The paper was finalised after review and suggested updates and comments from ID, UG, LS, and KW.

4.2.1. Paper 1 coversheet

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1703768	**Title**	Ms
First Name(s)	Maud Emma Louise		
Surname/Family Name	Teoh		
Thesis Title	Real-world effectiveness of oral anticoagulants in the prevention of stroke: emulation and extension of the ARISTOTLE trial using UK EHRs		
Primary Supervisor	Kevin Wing		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	BMJ Open		
When was the work published?	15 April 2021		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	NA		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I drafted the paper providing a summary of the planned analyses, I wrote the ISAC protocol this paper was based on and a more detailed Statistical Analysis Plan submitted to Bristol-Myers Squibb in the application for the individual patient trial data. Kevin Wing (KW) provided an example protocol previously published in the BMJ Open based on replication of a COPD trial [35]; his work helped inform the motivation for this PhD along with guiding the high-level plan of this trial replication work and served as an example on what details to include in the protocol paper. Ian Douglas (ID), Usha Gungabissoon (UG) and KW provided guidance on the data sources, methods, and the strengths and limitations of the proposed study. I performed the feasibility work involved in developing the protocol including review of the RCT protocol, development of appropriate codelists to apply the trial eligibility criteria to the EHR databases, and application of trial criteria to a sample of CPRD Gold data to estimate the cohort sample size. ID, UG, and KW advised on the implementation of the trial eligibility criteria. Liam Smeeth (LS) provided guidance in the development of some of the codelists in particular in terms of which recorded symptoms and diagnoses a GP may consider clinically relevant in assessing a patient's risk of bleeding. The paper was finalised after review and suggested updates and comments from ID, UG, LS, and KW.

SECTION E

Student Signature	[redacted]
Date	12 April 2024

Supervisor Signature	[redacted]
Date	12 April 2024

BMJ Open

Real-world effects of medications for stroke prevention in atrial fibrillation: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results

Emma Maud Powell ,[1] Ian J Douglas,[1] Usha Gungabissoon ,[2] Liam Smeeth,[1] Kevin Wing[1]

To cite: Powell EM, Douglas IJ, Gungabissoon U, *et al*. Real-world effects of medications for stroke prevention in atrial fibrillation: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results. *BMJ Open* 2021;**11**:e042947. doi:10.1136/bmjopen-2020-042947

► Prepublication history and additional material for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-042947).

Received 20 July 2020
Revised 26 November 2020
Accepted 23 February 2021

[1]Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK
[2]Epidemiology (Value Evidence and Outcomes), GSK, London, UK

Correspondence to
Emma Maud Powell;
maud.teoh@lshtm.ac.uk

ABSTRACT

Introduction Patients with atrial fibrillation experience an irregular heart rate and have an increased risk of stroke; prophylactic treatment with anticoagulation medication reduces this risk. Direct-acting oral anticoagulants (DOACs) have been approved providing an alternative to vitamin K antagonists such as warfarin. There is interest from regulatory bodies on the effectiveness of medications in routine clinical practice; however, uncertainty remains regarding the suitability of non-interventional data for answering questions on drug effectiveness and on the most suitable methods to be used. In this study, we will use data from Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (ARISTOTLE)—the pivotal trial for the DOAC apixaban—to validate non-interventional methods for assessing treatment effectiveness of anticoagulants. These methods could then be applied to analyse treatment effectiveness in people excluded from or under-represented in ARISTOTLE.

Methods and analysis Patient characteristics from ARISTOTLE will be used to select a cohort of patients with similar baseline characteristics from two UK electronic health record (EHR) databases, Clinical Practice Research Datalink Gold and Aurum (between 1 January 2013 and 31 July 2019). Methods such as propensity score matching and coarsened exact matching will be explored in matching between EHR treatment groups to determine the optimal method of obtaining a balanced cohort.

Absolute and relative risk of outcomes in the EHR trial-analogous cohort will be calculated and compared with the ARISTOTLE results; if results are deemed compatible the methods used for matching EHR treatment groups can then be used to examine drug effectiveness over a longer duration of exposure and in special patient groups of interest not studied in the trial.

Ethics and dissemination The study has been approved by the Independent Scientific Advisory Committee of the UK Medicines and Healthcare Products Regulatory Agency. Results will be disseminated in scientific publications and at relevant conferences.

Strengths and limitations of this study

- Selection of electronic health record patients matched to the randomised controlled trial (RCT) patients allows assessment of the ability of non-interventional methods to detect effectiveness of treatments for stroke prevention in atrial fibrillation (AF) within an RCT-analogous population.
- Combined Clinical Practice Research Datalink (CPRD) Gold and Aurum population broadly representative of the patients prescribed apixaban and warfarin for AF in routine clinical practice in the UK.
- Some of the criteria that were assessed for ARISTOTLE eligibility may not be well recorded in CPRD.
- Adherence to medication will need to be assessed based on proxy variables (time covered by prescription for the direct-acting oral anticoagulants, time in therapeutic range based on international normalised ratio measurements for warfarin); the different nature of these proxy variables means the adherence estimates may not be comparable.
- Ascertainment of outcomes via CPRD is based on recording as part of routine clinical care rather than for specifically detecting study outcomes.

INTRODUCTION

Background and rationale

Atrial fibrillation (AF) is a common cause of cardiac arrhythmia with symptoms including palpitations, fainting and shortness of breath; however, some patients may be asymptomatic. The prevalence of AF in the UK is estimated to be around 3%,[1] increasing from 0.2% in

people aged 45–54 years to 8.0% in those 75 and older.[2] The lack of organised atrial contraction in AF can lead to the formation of thrombi, meaning that patients with AF have a fivefold higher risk of stroke which is an important cause of morbidity and mortality.[3–5]

Current UK guidelines recommend use of prophylactic treatment with anticoagulation medication to reduce the risk of stroke. Warfarin, a vitamin K antagonist (VKA) and the previous standard anticoagulant treatment, has many treatment and dietary interactions requiring frequent monitoring of a patient's international normalised ratio (INR), to maintain anticoagulant activity within a narrow range (2.0–3.0). Low levels put the patient at a higher risk of stroke while high levels lead to a higher risk of bleeding.[6] In 2011, the first direct-acting oral anticoagulant (DOAC) dabigatran was approved for the treatment of AF in the European Union (EU); it was anticipated to provide easier to manage long-term anticoagulation therapy for patients with AF given the complex safety profile of warfarin. ARISTOTLE, a pivotal randomised controlled trial (RCT) of the DOAC apixaban, demonstrated superiority over warfarin for both prevention of stroke and safety (major bleeding) among individuals with AF.[7]

The generalisability of the ARISTOTLE trial is limited by the strict eligibility criteria; evidence on apixaban's treatment effect is therefore lacking for patients who would not have met the eligibility criteria such as those at increased bleeding risk or with severe comorbid conditions. The regulatory environment now demands evidence of treatment effectiveness outside the confines of randomised trials.[8 9] Non-interventional data sources have the potential to overcome many of the RCT limitations given that they contain data for a wide spectrum of patients treated with the drug in routine care, including patients who would have been not eligible for trials. Data collected as part of routine patient care such as electronic health record (EHR) provide a valuable opportunity to obtain evidence on the effectiveness of apixaban in a routine care setting. A key problem with using these data is that the absence of randomisation leaves them highly susceptible to confounding making it difficult to have confidence in the results.

To address this lack of confidence, this study will apply innovative matching approaches to create a trial-analogous non-interventional cohort for analysis. Records from UK EHRs will be matched to ARISTOTLE patients before using methods for matching between treatment groups within the non-interventional EHR data, creating an EHR population similar to the trial population that is well balanced by treatment group. If successful, estimates of effectiveness and safety of apixaban obtained from analysis of this ARISTOTLE-analogous cohort should be comparable with the results from the ARISTOTLE trial. The non-interventional analysis methods used to obtain these results may then be used to reliably estimate effects in understudied AF patient groups.

AIMS AND OBJECTIVES

The aims of this study are (1) to measure the association between anticoagulation treatments for stroke prevention in AF and time to stroke, systemic embolism (SE), myocardial infarction (MI), major bleeding and mortality among an ARISTOTLE-analogous cohort of patients from UK EHRs, and (2) to develop a methodological framework with in-built validation for using observational EHRs to answer questions about DOAC risks and benefits in patients not included or under-represented in the RCTs.

The specific objectives are to:

Objective 1. Check comparability of EHR data and robustness of methods for measuring stroke prevention medication effectiveness in an ARISTOTLE-analogous cohort using data from EHR data and by comparing with ARISTOTLE results.

Objective 2. Extension of trial findings: measure treatment effects of apixaban in patient groups excluded from ARISTOTLE.

Objective 3. Comparative effectiveness: compare treatment effectiveness between multiple individual anticoagulants (warfarin, apixaban, rivaroxaban, dabigatran) in ARISTOTLE-eligible cohorts and in patient groups excluded from ARISTOTLE.

METHODS AND ANALYSIS

Figure 1 (figure adapted from a study in real-world effects of medications for chronic obstructive pulmonary disease[10]) provides an overview of the study, covering the objectives and data sources used, and how RCT data will be used in Objective 1 to validate methods for analysing effectiveness of treatments for stroke prevention in AF in non-interventional data. Should Objective 1 prove successful the validated methods will be applied to unanswered questions in Objectives 2 and 3.

Study design

We will use a retrospective cohort study design using longitudinal data to evaluate the effects of prescribing apixaban versus warfarin and then versus other DOACs for prevention of stroke and SE in AF on key effectiveness and safety outcomes using non-interventional primary care data.

Setting/data sources

Patient data used in this study will be obtained from several sources: primary care data on UK National Health Service (NHS) patients from Clinical Practice Research Datalink (CPRD) Gold and Aurum databases, additional data on hospital events and mortality on UK NHS patients with linked data from the Hospital Episodes Statistics (HES) and Office for National Statistics (ONS) databases, and results from the ARISTOTLE trial.

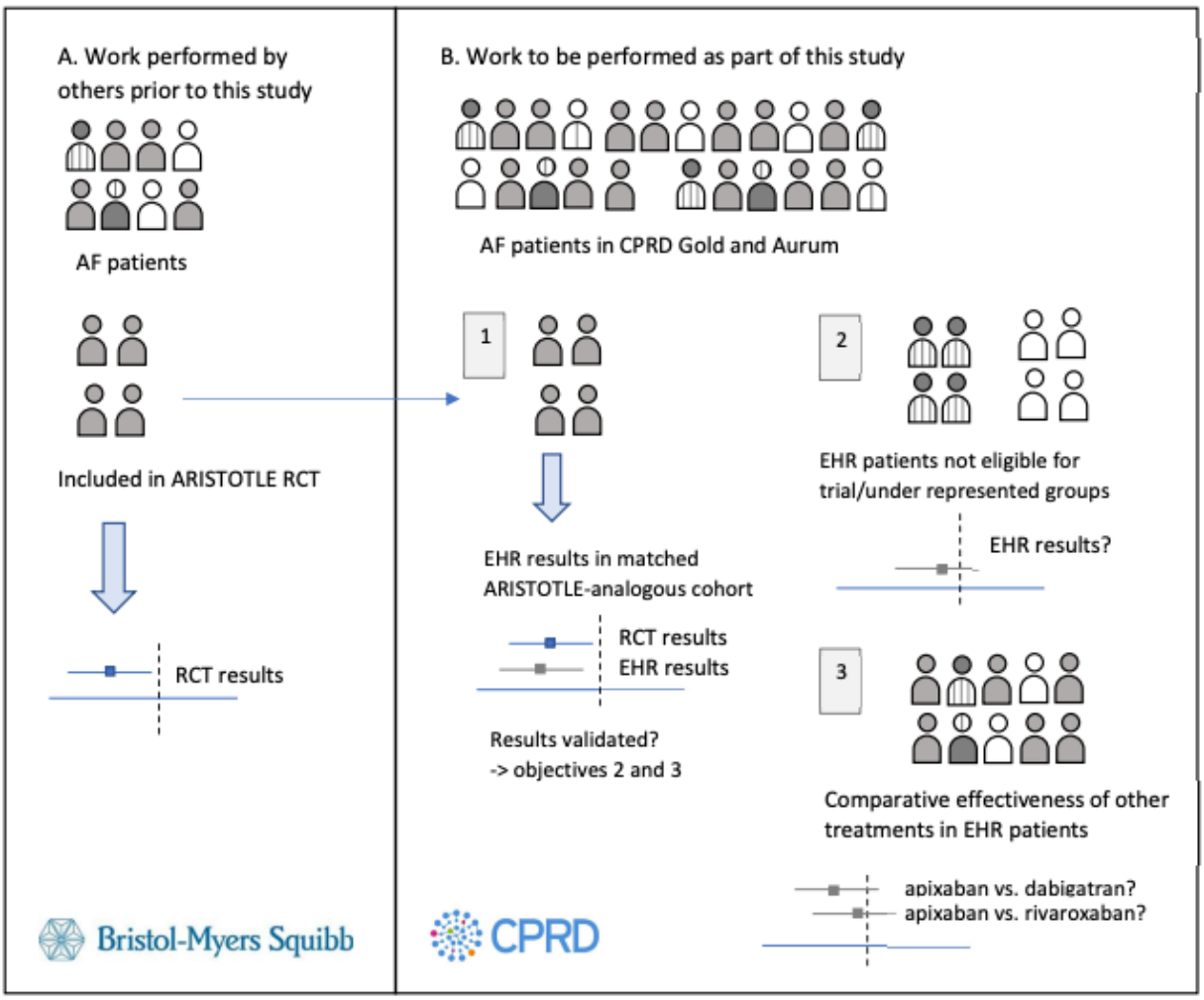


Figure 1 Overview of study objectives and sources of data for the real-world effects of medications for stroke prevention in AF study. AF, atrial fibrillation; CPRD, Clinical Practice Research Datalink; EHR, electronic health record; RCT, randomised controlled trial. (A) Work performed by others prior to this study. ARISTOTLE: RCT that investigated effectiveness and safety of apixaban vs warfarin in prevention of stroke and systemic embolism in AF patients. RCTs results inform clinical practice despite only a subset (based on trial inclusion and exclusion criteria) of the total population of AF patients being included in the RCTs of stroke prophylaxis treatments. (B) Work to be performed as part of this study. (1) Objective 1. A cohort of ARISTOTLE-analogous patients will be selected from UK EHRs (CPRD Gold and Aurum), by matching EHR patients prescribed apixaban to the apixaban patients included in the trial on baseline characteristics. EHR patients prescribed warfarin will then be matched to the trial-analogous EHR apixaban patients. An analysis of the effectiveness of apixaban vs. warfarin on prevention of stroke/systemic embolism will then be performed on this ARISTOTLE-analogous EHR cohort. If the results obtained are comparableto those obtained in ARISTOTLE, this will serve as a validation step, showing that data from the non-interventional CPRD Gold and Aurum sources can reliably be used to study stroke prevention treatment effects in AF. (2) Objective 2. The validated analysis techniques used for Objective 1 will then be used to study UK EHR patients who would not have been eligible for inclusion in an RCT or are under-represented in RCTs due to their age or presence of other comorbidities, for whom the comparative effects of anticoagulants in stroke prevention in AF is unclear.(3) Objective 3. The validated analysis techniques used for Objective 1 will then be used to compare effectiveness of apixaban vs warfarin, apixaban vsrivaroxaban and apixaban vs dabigatran.

ARISTOTLE

ARISTOTLE was a randomised, double-blind trial completed in 2011, comparing apixaban with warfarin in the prevention of stroke and SE. The trial included 18 201 patients with AF and at least one additional risk factor for stroke. The trial was designed to test for non-inferiority of apixaban compared with warfarin, and showed apixaban superiority for (1) the primary outcome of stroke or SE (HR 0.79, 95% CI 0.66 to 0.95),[7] (2) the safety endpoint of major bleeding (HR 0.69, 95% CI 0.60 to 0.80), and (3) death from any cause (HR 0.89, 95% CI 0.80 to 0.99). The ARISTOTLE findings led to the National Institute for Health and Care Excellence (NICE) guidelines on stroke prophylaxis in patients with AF recommending apixaban as a treatment. Baseline patient characteristics

from ARISTOTLE will be used in selection of participants in Objective 1.

CPRD Gold

CPRD Gold is a database containing anonymised data from over 625 primary care practices across the UK (approximately 13 million patient records) and is representative of the UK population with respect to age, gender and ethnicity.[11] Gold contains information on clinical diagnoses, prescribing, referrals, tests and demographic/lifestyle factors. General practices must meet prespecified standards for research-quality data to contribute data.

CPRD Aurum

CPRD Aurum contains primary care records similar to Gold but based on practices using EMIS software, whereas Gold has data from practices using Vision software. CPRD Aurum contains data on 19 million patients from 738 practices (10% of English practices) with 7 million active patients.[12]

Selection of participants

Participants will be selected from CPRD Gold and Aurum between 1 January 2013 and 31 July 2019. All patients will need to have been registered with a practice contributing research quality data for at least 6 months. Participant selection criteria will then vary by objective as detailed below.

Objective 1

An overview of each of the steps for participant selection for Objective 1 is provided in figure 2.

Step 1

We will select all (HES and ONS linked) patients in the EHR cohort (CPRD Gold and Aurum) who would have met the following *inclusion* criteria for the ARISTOTLE study, at least 6 months after patient registration in the database on or prior to the index date:

- Diagnosis of AF.
- Age 18+ years.
- One or more stroke risk factors (age 75 years or older; prior stroke, transient ischaemic attack or SE; congestive heart failure; diabetes mellitus; hypertension).

In ARISTOTLE, patients randomised to apixaban were new users of apixaban while both treatment arms were allowed to be previous users of warfarin, with patients stratified by prior warfarin/VKA exposure. To mirror ARISTOTLE, we will assess trial criteria for apixaban

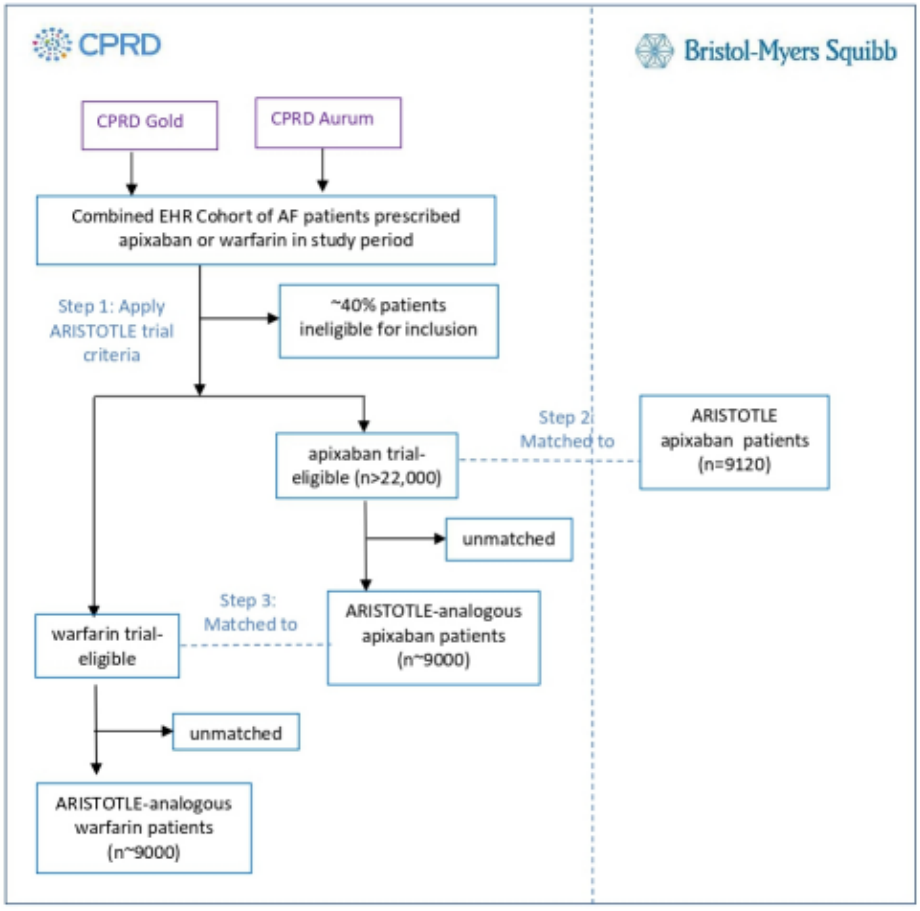


Figure 2 Flow chart illustrating the assembly of a matched trial-analogous cohort of EHR patients. AF, atrial fibrillation; CPRD, Clinical Practice Research Datalink; EHR, electronic health record.

patients on the date of their first prescription of apixaban while allowing patients prescribed warfarin to become eligible at any warfarin prescription date during the study period; furthermore, we will match ARISTOTLE in the proportion of new versus prevalent users in both treatment arms. We will then exclude patients who meet any of the following ARISTOTLE study *exclusion* criteria prior to their eligible-for-inclusion date:

- AF due to reversible causes.
- Mitral stenosis.
- Increased bleeding risk.
- Conditions other than AF requiring chronic anticoagulation.
- Persistent, uncontrolled hypertension.
- Active infective endocarditis.
- Current treatment with aspirin >165 mg/day.
- Simultaneous current treatment with both aspirin and a thienopyridine.
- Conditions likely to interfere with participation in the trial or cause death within 1 year.
- Recent alcohol or drug abuse, or psychosocial reasons making study participation impractical.
- Recent ischaemic stroke (within 7 days).
- Severe renal insufficiency.
- Alanine aminotransferase or aspartate aminotransferase >2× upper limit of normal (ULN) or total bilirubin ≥1.5× ULN.
- Platelet count ≤100x10^9/L
- Haemoglobin <90 g/L.
- Pregnancy or breast feeding.

Feasibility counts in Gold found approximately 60% of patients with AF prescribed apixaban met the ARISTOTLE trial criteria. Details of the algorithms used in applying the trial criteria to the EHR data are given in the online supplemental file.

Step 2

We will select a subset of apixaban patients from our EHR pool to create a cohort that matches the ARISTOTLE apixaban participants on a selection of the following baseline characteristics:

- Age.
- Sex.
- Body mass index (BMI).
- Systolic blood pressure (SBP).
- Congestive heart failure or left ventricular systolic dysfunction.
- Hypertension requiring treatment.
- Diabetes mellitus.
- Prior stroke/thromboembolism.
- Smoking status.
- Alcohol consumption.
- Level of renal impairment.
- Prior VKA/warfarin exposure.
- Labile INR in prior users of warfarin.
- Concomitant use of: aspirin, antiplatelet or non-steroidal anti-inflammatory drug, lipid-lowering drug therapy, or CYP3A4 inhibitor.

This step will generate a group of ARISTOTLE-analogous apixaban patients, with similar baseline characteristics to ARISTOTLE subjects at the point of randomisation (n~9000).

The variables selected are expected to influence the likelihood of the outcomes of interest. Exact selection of matching variables will depend on the quality and completeness of the data available and a balance will be struck between matched sample size and balance. Different methods to facilitate selection of a matched cohort will be explored, such as propensity score matching (PSM) and coarsened exact matching (CEM),[13] a non-parametric method that may give estimates with lower variance and bias for a given sample size compared than other methods.[14]

Step 3

The resulting trial matched sample of EHR apixaban patients will be matched to the warfarin ARISTOTLE-eligible EHR patients (figure 2) using a matching method such as PSM or CEM (final method selected based on giving optimal sample size vs balance). Risk set sampling will be employed in order to ensure similar duration of prior VKA/warfarin exposure for the prevalent users in the apixaban and warfarin EHR cohorts. The covariates for consideration in matching between EHR treatment arms or construction of a propensity score (PS) model will include the variables listed in step 2 along with additional EHR variables such as data source (Gold or Aurum), socioeconomic status and comorbidities. Each apixaban patient from the ARISTOTLE-eligible EHR patients will be matched 1:1 with the warfarin EHR patient with the closest match giving a trial-analogous cohort of ~18 000.

Step 4

The absolute rates and HR for the outcomes of interest (time to: stroke/SE, MI, major bleeding and mortality) will then be calculated. For the primary outcome (time to stroke/SE) the EHR results will be validated against the ARISTOTLE trial results using the criteria detailed in the Statistical Analysis section (Validation of observational results against ARISTOTLE data).

Objective 2

We will select patient groups who would not have been included in ARISTOTLE (and therefore would not have been included in the Objective 1 cohort) or who are under-represented in ARISTOTLE. Specifically, this will include patient groups such as patients with an AF diagnosis in the EHR cohort meeting these additional criteria:

- Severe comorbid condition: disease with a likelihood of causing death within 1 year or reasons making participation unpractical (such as dementia).

When matching the apixaban and warfarin patients within the patient groups for this objective, additional baseline variables will be considered compared with the list specified for Objective 1, Step 2; namely the H, A, and B components of the HAS-BLED score (Hypertension, Abnormal renal/liver

function, Stroke, Bleeding history or predisposition, Labile INR, Elderly (>65 years), Drugs/alcohol concomitantly) not included for Objective 1 matching due to being ARISTOTLE exclusion criteria. In these special patient populations the same outcomes as Objective 1 will be assessed, with absolute and relative rates calculated separately in each special patient group.

Objective 3

We will select all patients with AF who have a prescription for apixaban, warfarin, rivaroxaban or dabigatran in the treatment period (between 1 January 2013 and 31 July 2019). The ARISTOTLE trial criteria will be applied, followed by matching the warfarin, rivaroxaban and dabigatran ARISTOTLE-eligible EHR patients in turn to the trial-eligible EHR apixaban patients following the methodology outlined in Objective 1, Step 3. This process will result in the creation of three trial-eligible EHR cohorts: warfarin users matched to apixaban users, rivaroxaban users matched to apixaban users and dabigatran users matched to apixaban users. Matched cohorts of excluded patient groups will also be constructed to enable pairwise comparisons of treatment effects in these groups using the method outlined in Objective 2. In all cohorts, the same outcomes as Objective 1 will be assessed with both absolute and relative treatment effects compared.

Exposures, outcomes and covariates

Exposures

For all objectives, exposures will be determined using CPRD Gold and Aurum prescribing records and code lists for anticoagulant treatments with no restrictions placed on the dose prescribed.

For Objectives 1 and 2, use of apixaban is the primary exposure of interest and will be compared with warfarin.

For Objective 3, other stroke prevention treatments for AF will also be compared, namely dabigatran and rivaroxaban.

Outcomes

Outcomes to be measured are as follows:

- Stroke (ischaemic or haemorrhagic) or SE.
- Major bleeding.
- MI.
- All-cause mortality.
- Time to AF treatment change.

Outcomes will be ascertained using a combination of CPRD, HES and ONS data.

Covariates

The variables to be considered for matching patients are detailed in the selection of participants for Objective 1 (Step 2).

Sample size

Objective 1

ARISTOTLE included 9120 patients in the apixaban arm, therefore it was estimated a minimum of 15 000 EHR apixaban patients were needed for matching to be feasible. In CPRD Gold, approximately 8400 patients were eligible (January 2018). Aurum (June 2019) contained 23 526 apixaban patients with AF not registered in practices that had previously contributed data to Gold. Assuming the proportion of Aurum patients meeting ARISTOTLE eligibility criteria would be similar to the proportion in Gold (~60%) gave an estimate of 14 115 trial-eligible apixaban patients. Combining Gold and Aurum is therefore estimated to give >22 000 unique trial-eligible EHR apixaban patients.

Objectives 2 and 3

From feasibility counts, we are confident we will have sufficient numbers of patients to allow well-powered analyses for Objectives 2 and 3. For example, we estimate the number of people with no evidence of at least one additional risk factor for stroke for Objective 2 would be >3000 people in each exposure group.

Statistical analysis

Methods of analysis

ARISTOTLE used an intent-to-treat (ITT) approach for the primary efficacy analysis and an on-treatment approach for sensitivity analysis and safety outcomes. We will perform equivalent analyses by using two different censoring schemes: a primary censoring scheme censoring 5 years after index date (reflecting the maximum possible follow-up in ARISTOTLE) for the primary effectiveness analyses, and an on-treatment scheme censoring around time of last study drug for the sensitivity analysis and safety outcome. For the on-treatment censoring scheme, date of last exposure will be estimated using patient prescription data—to allow for drug half life, stockpiling of tablets and less than 100% adherence we will add 30 days after the apparent end of treatment.

Demographic and baseline variables will be presented before and after matching steps. As the primary analysis accounts neither for treatment switching nor discontinuation, the proportion of patients discontinuing treatment and time to treatment discontinuation will be tabulated.

The primary effectiveness endpoint is time to first occurrence of confirmed stroke (ischaemic, haemorrhagic or unspecified type) or SE during the study, regardless of whether the subject is receiving treatment at the time (primary censoring scheme). Comparisons will be made according to prescribed treatment (apixaban vs warfarin).

All time to event endpoints will be analysed using a Cox proportional hazards model including treatment group as a covariate and prior warfarin/VKA status (experienced, naïve). Point estimates and two-sided 95% CIs will be constructed for the outcome. Absolute event rates of all outcomes of interest will also be calculated.

Secondary outcomes cover the key safety outcome of major bleeding and the individual outcomes of stroke, SE, MI and mortality. Secondary outcomes other than major bleeding will use the ITT censoring scheme, major bleeding will use the on-treatment censoring scheme.

Validation of observational results against ARISTOTLE data

In Objective 1 alone, we will validate the findings from our primary analysis against ARISTOTLE by determining

whether results are compatible with the trial results. ARISTOTLE demonstrated superiority of apixaban over warfarin for the primary endpoint (HR 0.79, 95% CI 0.66 to 0.95).[7] The treatment effect seen with EHR data may be weaker than that seen in ARISTOTLE.

An analysis of EU patients in ARISTOTLE showed a smaller treatment difference for the primary endpoint and death: HR for stroke/SE 0.92 (95% CI 0.56 to 1.52), all-cause death 0.89 (95% CI 0.68 to 1.18). The European Medicines Agency Assessment Report suggested the smaller treatment effect may have been due to superior INR control in the warfarin arm of the EU subgroup (median time in therapeutic range (TTR) 68.93%)[15]; this study could provide additional evidence on this point.

Either a result of superiority or non-inferiority will be considered compatible with ARISTOTLE results. We have set two criteria that must be met to conclude results are consistent with the trial result:

1. The effect size must be clinically comparable with the ARISTOTLE findings; the HR for time to stroke/SE with the EHR must be between 0.69 and 0.99. This range is not symmetrical around the ARISTOTLE estimate of 0.79 as it is anticipated the treatment effect in routine clinical care may be weaker than that seen in the optimised setting of a clinical trial.
2. The upper limit of the 95% CI for the rate ratio must be less than 1.52 (upper limit in the EU subgroup of ARISTOTLE).

In addition, if the upper limit of the 95% CI is less than 1 then superiority of apixaban versus warfarin will be concluded.

In order to understand the extent to which the EHR population resembles the ARISTOTLE trial population the absolute event rates of the outcomes will be compared.

Sensitivity analyses

Primary and secondary effectiveness outcomes will also be analysed using the on-treatment censoring scheme to investigate whether the extent of treatment discontinuation compromises confidence in the effectiveness analyses.

Exclusion of patient time post-treatment discontinuation in the safety and sensitivity analyses might bias results towards a conclusion of no difference[16] and risks selection bias due to attrition[17]; the set of patients who switch or discontinue treatment will therefore be examined to ascertain whether biases of this nature may have occurred.

Additional analyses may be performed using methods such as inverse probability of censoring weighting (IPW) or a rank-preserving structural failure time model to estimate the treatment effect that would have been observed in the absence of treatment switching. We will explore the impact of time-varying eligibility by using methods such as a modified treatment strategy IPW.[17]

Adherence will be estimated in the EHR cohort to enable comparisons with the trial and investigate the extent to which this may have influenced differences in treatment effect observed. For apixaban, we will calculate the proportion of days covered (PDC) over a patient's time when on treatment as a measure of adherence. Warfarin dose is poorly recorded in EHR, therefore warfarin adherence will be estimated by looking at adherence to other long-term daily medications as a proxy measure and by looking at INR control by calculating per cent INR TTR as a measure of overall warfarin treatment regime adherence.

We will perform a supplementary analysis in patients deemed adherent (PDC ≥80% matching ARISTOTLE compliance limit) along with an exploratory subgroup analysis by INR TTR. The different nature of the proxy variables used for adherence in the DOACs (PDC) compared with warfarin (INR TTR) means that the adherence estimates may not be comparable; should great differences in adherence be observed between these exposure groups the definitions of adherence used may need to be reassessed.

Apixaban was a newly available drug with a low number of patients having a prescription in the first year it was available[18]; we will therefore perform a sensitivity analysis with the start of the study period shifted forwards a year to January 2014 to investigate the impact of inclusion of early adopters who may differ from later adopters of a new drug.

Plan for addressing confounding

In the study period, apixaban was a newly available treatment leading to the possibility of channelling bias. For Objective 1, by applying trial eligibility criteria to both treatment cohorts and matching using the baseline covariates we should avoid channelling bias. To handle confounding, treatment arms will be matched using the optimal method selected. Unmeasured or unknown confounding may remain and this will be explored in the analysis and discussion of results.

Missing baseline data

UK EHR data have been shown to be almost complete for drug prescribing and information on important comorbidity is well recorded. For some variables such as renal function and alcohol intake, a patient is more likely to have no data entered if there is no overt clinical evidence of abnormality; in such cases, we may take a pragmatic approach categorising into a parameter ('evidence of' vs 'no evidence of') with those with no data included in the 'no evidence of' group. For BMI and SBP, we cannot assume data are missing at random as we expect a patient is less likely to have these recorded if they appear at a healthy weight and do not have hypertension, respectively, or if they have a lower comorbidity burden. Furthermore, as the proportion of patients with missing baseline BMI or SBP is expected to be low (approximately 4% for BMI and <1% for SBP[18]), these patients will be excluded from the trial-eligible cohort.

Missing prescription data

Treatment may be initiated in secondary care, meaning the first prescription of patients newly initiating treatment or switching treatments is missing; to account for this we will perform a sensitivity analysis where those newly initiating treatment are assigned an earlier derived index date. Hospitalised patients may have prescriptions in secondary care leading to treatment gaps in their primary care data.

We will investigate the occurrence of hospitalisation around treatment discontinuation and assess the potential impact on the results of missed events by performing a sensitivity analysis with different extended derived dates of last dose. Some concomitant drugs used in determining eligibility and matching patients are available over the counter (OTC), meaning we may miss that patients are exposed to these; we expect OTC use of these drugs to be similar in both treatment groups.

Missing outcome data

EHR data are shown to be almost complete for mortality.[19] Patient deaths missing from EHRs are expected to be missing at random equally in both treatment arms, thereby not altering the overall direction of treatment effect. The classification of unspecified stroke type will cause uncertainty in the main safety endpoint and may lead to a lower event rate for major bleeding compared with the trial; this would affect the power but should not affect the treatment effect seen as events are expected to be missing at random from both treatment arms.

Limitations of the study design, data sources and analytical methods

Some of the criteria assessed for ARISTOTLE eligibility may not be well recorded in CPRD, criteria such as alcohol and drug abuse may not be captured for all patients. For criteria such as 'increased bleeding risk', it is unclear which codes to include and timescale to consider. These limitations are consistent with our aim to select a population as similar as possible to ARISTOTLE while acknowledging differences will remain. The most important risk factors for the primary outcome of stroke (the components of CHA2DS2-VASc score for AF stroke risk) are mostly well recorded in CPRD.[20]

There are differences in the coding systems used by the two EHR data sources and completeness of coding may differ between the two; the potential impact of this will be ascertained by comparisons of rates of diagnoses, baseline variables and prescriptions of interest. Inclusion of data source as a matching variable should prevent discrepancy between the sources from biasing results. We will explore different methods of combining Gold and Aurum, namely analysing separately by database and combining the results as a meta-analysis as an alternative to combining data before analysis.

The main focus of the study is validation of our methodology through assembling a cohort of patients comparable to the patients in ARISTOTLE and finding similar results to the trial. Criteria to determine the success of the methodology have been prespecified in the protocol. Given the use of CPRD data to determine treatment effectiveness is not yet well established, a finding that these data are not suitable to answer questions on intended effectiveness will be a useful conclusion.

Patient and public involvement

No patient was involved.

ETHICS AND DISSEMINATION

Approval by ethics and scientific committees

An application for scientific approval related to use of CPRD data was approved by the Independent Scientific Advisory Committee of the Medicines and Healthcare Products Regulatory Agency (MHRA).

Dissemination plans

The results of the study will be submitted to peer-reviewed journals and presented at conferences. Relevant charities will be contacted for guidance on dissemination of results to patients in an accessible manner. We will communicate with NICE to convey any results relevant to the guidance they have issued on AF, and with the MHRA if findings may impact the risk/benefit profile of anticoagulation treatments in patients with AF.

Contributors EMP, KW, IJD, UG and LS contributed to study question and design. EMP wrote the first draft of the protocol manuscript (based on the original proposal to MRC, ISAC that EMP, KW, IJD, UG and LS all contributed to). EMP, KW, IJD, UG and LS contributed to further drafts and approved the final version.

Funding This work was supported by the Medical Research Council through an MRC LID studentship (grant number MR/N013638/1).

Competing interests IJD reports grants from GlaxoSmithKline, NIHR, ABPI and MRC and holds stock in GlaxoSmithKline. LS reports grants from Wellcome, MRC, NIHR, BHF, Diabetes UK, ESRC and the EU; grants and personal fees for advisory work from GSK; and historical personal fees for advisory work from AstraZeneca. He is a Trustee of the British Heart Foundation. UG is an employee of and holds shares in GlaxoSmithKline.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs

Emma Maud Powell http://orcid.org/0000-0001-9427-9468
Usha Gungabissoon http://orcid.org/0000-0002-2040-1763

REFERENCES

1 Adderley NJ, Ryan R, Nirantharakumar K, *et al*. Prevalence and treatment of atrial fibrillation in UK general practice from 2000 to 2016. *Heart* 2019;105:27–33.
2 Davis RC, Hobbs FDR, Kenkre JE, *et al*. Prevalence of atrial fibrillation in the general population and in high-risk groups: the echoes study. *Europace* 2012;14:1553–9.
3 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. *Stroke* 1991;22:983–8.
4 Brønnum-Hansen H, Davidsen M, Thorvaldsen P, *et al*. Long-Term survival and causes of death after stroke. *Stroke* 2001;32:2131–6.
5 Adamson J, Beswick A, Ebrahim S. Is stroke the most common cause of disability? *J Stroke Cerebrovasc Dis* 2004;13:171–7.
6 Nelson WW, Wang L, Baser O, *et al*. Out-of-range Inr values and outcomes among new warfarin patients with non-valvular atrial fibrillation. *Int J Clin Pharm* 2015;37:53–9.

7 Granger CB, Alexander JH, McMurray JJV, *et al*. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011;365:981–92.
8 U.S. Department of Health and Human Services F and D. Guidance for Industry Postmarketing Studies and Clinical Trials — Implementation of Section 505(o)(3) of the Federal Food, Drug, and Cosmetic Act
9 European parliament and the council of the European union. *Regulation (EU) NO 1235/2010 of the European Parliament and of the Council of 15 December*, 2010.
10 Wing K, Williamson E, Carpenter JR, *et al*. Real-World effects of medications for chronic obstructive pulmonary disease: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results. *BMJ Open* 2018;8:e019475.
11 Herrett E, Gallagher AM, Bhaskaran K, *et al*. Data resource profile: clinical practice research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
12 Wolf A, Dedman D, Campbell J, *et al*. Data resource profile: clinical practice research Datalink (CPRD) aurum. *Int J Epidemiol* 2019;48:1740–1740g.
13 Iacus SM, King G, Porro G. Causal inference without balance checking: Coarsened exact matching. *Polit. anal.* 2012;20:1–24.
14 King G, Nielsen R, Coberley C. Comparative effectiveness of matching methods for causal inference. Available: https://gking.harvard.edu/publications/comparative-effectiveness-matching-methods-causal-inference [Accessed 4 Sep 2018].
15 European Medicines Agency. *Committee for medicinal products for human use (CHMP). assessment report Eliquis apixaban. Procedure No.:EMEA/H/C/002148/X/04/G.EMA/641505/2012*, 2012.
16 Jones B, Jarvis P, Lewis JA, *et al*. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996;313:36–9.
17 Schnitzer ME, Platt RW, Durand M. A tutorial on dealing with time-varying eligibility for treatment: comparing the risk of major bleeding with direct-acting oral anticoagulants vs warfarin. *Stat Med* 2020;39:4538–50.
18 Vinogradova Y, Coupland C, Hill T, *et al*. Risks and benefits of direct oral anticoagulants versus warfarin in a real world setting: cohort study in primary care. *BMJ* 2018;362:k2505.
19 Gallagher AM, Williams T, Leufkens HGM, *et al*. The impact of the choice of data source in record linkage studies estimating mortality in venous thromboembolism. *PLoS One* 2016;11:e0148349.
20 Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the general practice research database: a systematic review. *Br J Gen Pract* 2010;60:e128–36.

4.3. Additional detail on methods outlined in the published protocol

4.3.1. Additional information on data sources for the emulation of ARISTOTLE

Based on the feasibility work described in Chapter 2 it was initially planned to use both CPRD Aurum and Gold datasets combined. Applying the ARISTOTLE criteria to CPRD Aurum resulted in a sample size in Aurum alone more than sufficient to enable selection of a subset of patients matching ARISTOTLE. CPRD Gold and Aurum have differences in the way data are recorded meaning combining both datasets would introduce more variability into the data compared with restricting the EHR cohort to come from only 1 data source. Furthermore, with many GP practices in England and Wales having transitioned from the software system used for Gold to the system used for Aurum, the CPRD Aurum database had a larger volume of more recent data.

4.3.2. Additional information on codelists

A total of 55 codelists (Table 4.1) were developed for trial eligibility criteria, outcomes of interest, and covariates of interest. Code lists were developed by searching for and reviewing pre-existing Read code, medication, and ICD-10 codelists and using or modifying as appropriate. Suitable Read code codelists were mapped to equivalent Snomed codelists in a systematic way by mapping between the codes, matching on textual terms, identifying all Snomed terms related to any unique concept IDs captured in these earlier steps, and a final additional search for matching terms. Codelists were reviewed to ensure the terms identified were suitable to the intended criteria, outcome, or covariate and clinical input was obtained on any codelists not well supported in past literature. The codelists used are presented in the appendix. Derived datasets were created by extracting records matching the codelists and cleaning the data.

Table 4.1 List of Codelists Used

Codelist File Name	Description	Source
codelist_afib_aurum.txt	Snomed codes for Atrial Fibrillation for CPRD Aurum	Maud Teoh
codelist_oac_aurum.txt	Codes for Oral anticoagulant medications extracted from CPRD Aurum.	Maud Teoh
codelist_stroke_tia_se_aurum.txt	Snomed codes for Stroke, transient ischemic attack (TIA), systemic embolism (SE) for CPRD Aurum	Maud Teoh
codelist_stroke_tia_se_hes.txt	ICD-10 codes for stroke, transient ischemic attack (TIA), systemic embolism (SE) for HES	Maud Teoh
codelist_chf_lvef_aurum.txt	Snomed codes for congestive heart failure (CHF), reduced left ventricular ejection fraction (LVEF) for CPRD Aurum	Maud Teoh
codelist_chf_lvef_hes.txt	ICD-10 codes for congestive heart failure (CHF), reduced left ventricular ejection fraction (LVEF) for HES	Maud Teoh
codelist_diabetes_aurum.txt	Snomed codes for diabetes for CPRD Aurum	Maud Teoh
codelist_hypertension_aurum.txt	Snomed codes for hypertension for CPRD Aurum	Maud Teoh
codelist_antihypertensive_aurum.txt	Product codes for antihypertensive medications extracted from CPRD Aurum.	Maud Teoh
codelist_rev_afib_aurum.txt	Snomed codes for reversible atrial fibrillation for CPRD Aurum	Maud Teoh
codelist_mitral_stenosis_aurum.txt	Snomed codes for mitral stenosis for CPRD Aurum	Maud Teoh
codelist_bleed_risk_aurum.txt	Snomed codes for increased bleeding risk for CPRD Aurum	Maud Teoh
codelist_bleed_risk_hes.txt	ICD-10 codes for increased bleeding risk for HES	Maud Teoh
codelist_heart_valve_aurum.txt	Snomed codes for heart valve for CPRD Aurum	Maud Teoh
codelist_pe_dvt_aurum.txt	Snomed codes for pulmonary embolism (PE), deep vein thrombosis (DVT) for CPRD Aurum	Maud Teoh
codelist_endocarditis_aurum.txt	Snomed codes for endocarditis for CPRD Aurum	Maud Teoh
codelist_aspirin_aurum.txt	Product codes for aspirin extracted from CPRD Aurum	Maud Teoh
codelist_thienopyridine_aurum.txt	Product codes for thienopyridine extracted from CPRD Aurum	Maud Teoh
codelist_severe_comorbid_aurum.txt	Snomed codes for severe comorbid conditions for CPRD Aurum	Maud Teoh, Kevin Wing
codelist_alcohol_drug_abuse_aurum.txt	Snomed codes for alcohol or drug abuse for CPRD Aurum	Maud Teoh

codelist_renal_aurum.txt	Snomed codes for renal function for CPRD Aurum	Maud Teoh
codelist_renal_hes.txt	ICD-10 codes for severe renal disease and dialysis for HES	Maud Teoh
codelist_liver_aurum.txt	Snomed codes for liver function for CPRD Aurum	Maud Teoh
codelist_platelet_aurum.txt	Snomed codes for platelets for CPRD Aurum	Maud Teoh
codelist_hemoglobin_aurum.txt	Snomed codes for haemoglobin for CPRD Aurum	Maud Teoh
codelist_pregnant_breasteefing_aurum.txt	Snomed codes for pregnancy or breastfeeding for CPRD Aurum	Maud Teoh
codelist_blood_pressure_aurum.txt	Snomed codes for blood pressure for CPRD Aurum	Maud Teoh
codelist_ethnicity_aurum.txt	Snomed codes for ethnicity for CPRD Aurum	Maud Teoh
codelist_smoking_aurum.txt	Snomed codes for smoking for CPRD Aurum	Maud Teoh
codelist_alcohol_aurum.txt	Snomed codes for alcohol for CPRD Aurum	Maud Teoh
codelist_height_weight_bmi_aurum.txt	Snomed codes for height, weight, body mass index (BMI) for CPRD Aurum	Maud Teoh
codelist_amiodarone_aurum.txt	Product codes for amiodarone extracted from CPRD Aurum	Maud Teoh
codelist_digoxin_aurum.txt	Product codes for digoxin extracted from CPRD Aurum	Maud Teoh
codelist_statin_aurum.txt	Product codes for statins extracted from CPRD Aurum	Maud Teoh
codelist_nsaid_aurum.txt	Product codes for non-steroidal anti-inflammatory drug (nsaid) extracted from CPRD Aurum	Maud Teoh
codelist_antacid_aurum.txt	Product codes for antacids extracted from CPRD Aurum	Maud Teoh
codelist_ppi_aurum.txt	Product codes for proton pump inhibitors (PPI) extracted from CPRD Aurum	Maud Teoh
codelist_h2ra_aurum.txt	Product codes for H2 receptor antagonist (H2RA) extracted from CPRD Aurum	Maud Teoh
codelist_pad_aurum.txt	Snomed codes for peripheral artery disease (PAD) for CPRD Aurum	Maud Teoh
codelist_aortic_plaque_aurum.txt	Snomed codes for aortic plaque for CPRD Aurum	Maud Teoh
codelist_mi_aurum.txt	Snomed codes for myocardial infarction (MI) for CPRD Aurum	Maud Teoh

codelist_peptic_aurum.txt	Snomed codes for peptic ulcer for CPRD Aurum	Maud Teoh
OpenSAFELY codelist for lung cancer and OpenSAFELY codelist for cancer excluding lung and haematological	Snomed codes for cancer (excluding haematological) for CPRD Aurum https://www.opencodelists.org/codelist/opensafely/lung-cancer-snomed/2020-04-15/#full-list https://www.opencodelists.org/codelist/opensafely/cancer-excluding-lung-and-haematological-snomed/2020-04-15/	OpenSAFELY authors
codelist_connect_tissue_aurum.txt	Snomed codes for connective tissue disorder for CPRD Aurum	Maud Teoh
codelist_copd_aurum.txt	Snomed codes for chronic obstructive pulmonary disease (COPD) for CPRD Aurum	Maud Teoh
OpenSAFELY snomed codelist for haematological cancer	Snomed codes for haematological cancer for CPRD Aurum https://www.opencodelists.org/codelist/opensafely/haematological-cancer-snomed/2020-04-15/#full-list	OpenSAFELY authors
codelist_hemiplegia_aurum.txt	Snomed codes for hemiplegia for CPRD Aurum	Maud Teoh
codelist_aids_hiv_aurum.txt	Snomed codes for AIDS or HIV for CPRD Aurum	Maud Teoh
OpenSAFELY Codelist for Chronic Liver Disease (opensafely/chronic-liver-disease-snomed)	Snomed codes for liver disease for CPRD Aurum	OpenSAFELY authors
codelist_inr_aurum.txt	Snomed codes for international normalised ratio (INR) for CPRD Aurum	Maud Teoh
OpenSAFELY codelist for fall	Snomed codes for fall for CPRD Aurum	OpenSAFELY authors
codelist_stroke_embol_hes.txt	ICD-10 codes for stroke or systemic embolism for HES	Maud Teoh
codelist_major_bleed_hes.txt	ICD-10 codes for major bleeding for HES	Maud Teoh, Turki Bin Hammad
codelist_mi_hes.txt	ICD-10 codes for myocardial infarction (MI) for HES	Maud Teoh
codelist_pe_dvt_hes.txt	ICD-10 codes for pulmonary embolism (PE), deep vein thrombosis (DVT) for HES	Maud Teoh

Note: For ICD-10 codelists where a higher-level code is specified eg ‘I21’, all lower-level codes under this code are included (I21.0, I21.1,… etc).

4.3.3. Additional information on algorithms

Several covariates required the use of algorithms to aid classification, namely classification of ethnicity, smoking, and alcohol use; details on the algorithms used for these are in the appendix. There are a number of methods that can be used to measure quality of warfarin treatment using INR measurements, most commonly the proportion of INR measurements that are in optimal range over a certain time period and Rosendaal's method using linear interpolation to calculate the proportion of days a patient had INR values in optimal range. For this study I matched ARISTOTLE in using Rosendaal's method to calculate TTR.

4.3.3.1. Rosendaal's method of calculation for proportion of time in therapeutic range

Derivation of time in therapeutic range followed the method used in ARISTOTLE - Rosendaal's method of linear interpolation. In this method it is assumed that changes between consecutive INR measurements are linear over time. Using an optimal range for INR [2-3] the calculation works as follows:

Let INR_i be a patient's INR value recorded on visit i where $i = 1,2 \ldots,k-1$
Let INR_j be the next INR value recorded for a patient on visit $j = i+1$

1. Calculate the magnitude of the shift from 1 INR measurement to the next
   $shift$ = abs($INR_j - INRi$)
2. Calculate the proportion of this shift that was within the therapeutic INR range using Table 4.2
3. Estimate the number of days between the consecutive visits at which the INR measurements were obtained, $visit_i$ and $visit_j$ that were within therapeutic range:
   $number\ of\ days\ in\ range_{i,j} = proportion \times (day_j - day_i)$
4. Total Time in Therapeutic Range (TTR) is then the sum of the estimated number of days in range divided by the total number of days between the visits

$$TTR = \frac{\sum_{i=1}^{i=k-1} number\ of\ days\ in\ range_{i,i+1}}{\sum_{i=1}^{i=k} number\ of\ days_{i,k}}$$

Table 4.2 Proportion of shift in consecutive INR measurements within therapeutic range

INR_i	INR_j	Proportion of Shift Within Therapeutic Range
within [2-3]	within [2-3]	1
within [2-3]	> 3	$abs\ [(3 - INR_i)/(INR_j - INR_i)]$
within [2-3]	< 2	$abs\ [(INR_i - 2)/(INR_j - INR_i)]$
> 3	within [2-3]	$abs\ [(3 - INR_j)/(INR_j - INR_i)]$
> 3	> 3	0
> 3	< 2	$abs\ [1/(INR_j - INR_i)]$
< 2	within [2-3]	$abs\ [(INR_j - 2)/(INR_j - INR_i)]$
< 2	> 3	$abs\ [1/(INR_j - INR_i)]$
< 2	< 2	0

Note: Using a therapeutic range of INR [2.0 to 3.0] and assuming INR_i and INR_j are consecutive INR measurements for an example patient ie j=i+1.

To illustrate the calculation of INR consider the example of a patient with INR value of 1.8 recorded on a first visit, 2.8 recorded on their next visit 20 days later, and 3.2 recorded on a 3rd occasion 10 days after the 2nd measurement. In this example the magnitude of the shift between the 2 pairs of consecutive INR measurements is

$$shift_{1,2} = abs(INR_2 - INR_1) = abs(2.8 - 1.8) = 1.0$$

$$shift_{2,3} = abs(INR_3 - INR_2) = abs(3.2 - 2.8) = 0.4$$

The proportion of the shifts within the therapeutic INR range is then

$$proportion_{1,2} = abs\left(\frac{INR_2 - 2}{INR_2 - INR_1}\right) = abs\left(\frac{2.8 - 2}{1.0}\right) = 0.8$$

$$proportion_{2,3} = abs\left(\frac{3 - INR_2}{INR_3 - INR_2}\right) = abs\left(\frac{3 - 2.8}{0.4}\right) = 0.5$$

The number of days between these consecutive visits which are estimated to have been within therapeutic INR range is then

$$number\ of\ days\ in\ range_{1,2} = 0.8 \times 20 = 16\ days$$
$$number\ of\ days\ in\ range_{2,3} = 0.5 \times 10 = 5\ days$$

Giving a total TTR over the 3 visits of:

$$TTR = \frac{16 + 5}{30} = \frac{21}{30} = 0.7$$

In this example we see the hypothetical patient has a TTR of 0.7 demonstrating they are spending the majority of their time within the optimal INR range. By contrast, the proportion of INR measurements within therapeutic range for the same patient would be only 0.33 showing the impact of the choice of method to assess quality of INR control.

4.3.4. Additional information on step 2: selection of CPRD Aurum patients matching the ARISTOTLE participants

The original plan for this thesis involved the use of individual patient data from ARISTOTLE to enable matching of individual EHR patients 1:1 to each trial participant in the apixaban arm of ARISTOTLE. In the first year of the PhD I submitted an application (including a statistical analysis plan) to Bristol-Myers Squibb for use of the individual patient data in ARISTOTLE. The application was approved, however a subsequent review by Bristol-Myers Squibb found that the informed consent signed by patients in the trial prevented third party researchers using the data. Lack of access to this data necessitated a modification to the methods to match to ARISTOTLE using only publicly available information consisting of summary statistics of the baseline characteristics of the treatment arms and subgroup analyses published based on the trial. A search of publications on ARISTOTLE likely to be instructive in matching to the trial participants identified the sources of information listed in Table 4.3.

Table 4.3 Sources Used in Matching to ARISTOTLE

Source(s)	Description/ Use in this Study
ARISTOTLE protocol and SAP Bristol-Myers Squibb and Pfizer, Inc. [see Appendix]	Detailed the study design, derivation of non-inferiority margin, hypothesis and sample size calculation, inclusion and exclusion criteria, randomisation and stratification factors, outcome definition, and analysis methods of ARISTOTLE
NEJM key publication Apixaban versus Warfarin in Patients with Atrial Fibrillation. C B Granger et al. (42)	Summarised the key ARISTOTLE baseline characteristics and ARISTOTLE results
FDA NDA Review Rose M, Beasley N (131)	Summarised ARISTOTLE baseline characteristics and results in more detail than the NEJM publication
EMA NDA Review (132)	Summarised ARISTOTLE baseline characteristics and results in more detail than the NEJM publication and gave results in EU subset, context and questions on relevance to EU patients
NICE	Summarised ARISTOTLE baseline characteristics and results and gave context and questions on relevance to UK population
ARISTOTLE Outcomes by Sex. Described in a paper by D Vinereanu et al. (133)	Summarised ARISTOTLE baseline characteristics and results by sex
ARISTOTLE Outcomes by Age. Described in a paper by S Halvorsend et al. (134)	Summarised ARISTOTLE results by age and gave distribution of elderly age groups

4.3.4.1. Development of matching approach

Table 1 Baseline Characteristics of the Patients from the NEJM publication of the ARISTOTLE trial (42) was used as a starting point to represent the target population. The variables in the table were reviewed to consider which would be appropriate or possible to attempt to match in the CPRD Aurum cohort and which were the most important to match given the indication, treatments, and outcomes of interest to this study. The key variables to match that were available in CPRD Aurum were determined to be:

- Age
- Sex
- Prior use of VKA
- Qualifying stroke risk factors and $CHADS_2$ score
- Renal function

These variables are important predictors of the key outcomes of stroke, bleeding, and all-cause mortality. Variables that were partially matched or found to not require matching are summarised in Table 4.4 and characteristics from 'Table 1' of the ARISTOTLE publication that were not matched are detailed in Table 4.5.

Table 4.4 Variables from ARISTOTLE partially matched or did not require matching

Variable	Reason no matching required or partially matched
Systolic blood pressure	The matching method included hypertension and the ARISTOTLE eligibility criteria excluded those with uncontrolled hypertension. Distribution of SBP of the selected CPRD Aurum cohort matched the trial cohort without including this as a variable.
Weight	Though Table 1 in ARISTOTLE summarised weight, ARISTOTLE participants differed from CPRD Aurum patients in ethnicity making body mass index (BMI) a more appropriate measure than weight to match to the trial. A 'partial' matching approach was taken in which the probability of being sampled into the CPRD Aurum cohort was adjusted based on BMI category.
Prior myocardial infarction (MI)	After applying the trial eligibility criteria and matching on the other variables the selected CPRD Aurum cohort (12.3%) was similar to the trial (14.5%) so further matching was not performed.
Prior clinically relevant or spontaneous bleeding	After applying the trial eligibility criteria and matching on the other variables the selected CPRD Aurum cohort (17.3%) was similar to the trial (16.7%) so further matching was not performed.
History of fall within previous year	The prevalence of this variable depended on how it was defined (for example restricting to hospital records or including CPRD records, inclusion of proxy measures from HES such as broken bones or head injuries, how to handle CPRD records such as 'history of fall' where time frame is not clear) thus the EHR variable may not be directly comparable to a detailed question in an RCT screening or baseline assessment. By matching to the trial on age and matching on the proportion of patients in the age groups 75-80, 80+, 90+, the proportion of patients with a history of fall was likely to be similar.

Table 4.5 Variables from ARISTOTLE not matched

Variable	Reason not matched
Region	All patients in CPRD Aurum from the same region (Europe)
Race/ethnicity	Low number of patients from Asian and/or Hispanic ethnicity in CPRD Aurum (majority of patients of white ethnicity)
Type of atrial fibrillation	Type of AF not recorded in CPRD Aurum for the majority of patients
Medications at the time of randomisation	Choice of treatment for conditions such as hypertension and heart failure may differ across countries. Attempting to match to the trial on medications at baseline may therefore result in selection of patients in CPRD Aurum which are less likely to reflect typical treatment for patients in the UK.

To further characterise the trial population, a search for publications containing baseline characteristics from the trial found the FDA, EMA, and NICE reviews of the trials gave some additional information on the baseline characteristics. Subgroup studies using the trial data gave additional information for example analysis by sex and in the elderly.

4.3.4.2. Use of ARISTOTLE baseline characteristics to select patients in CPRD Aurum

The sources identified in Table 4.3 were used to define a target multivariable covariate distribution based on the pairwise and higher order observed covariate profile sub-distributions. For example, the number of participants with each stroke risk factor in the apixaban arm gave the following target distributions:

Congestive heart failure = C = 2784

Hypertension = H = 7962

Age ≥ 75 years = $A = 2850$

Diabetes = D = 2284

History of stroke or TIA = S approx. 1650

Combinations of stroke risk factors that can result in each $CHADS_2$ score along with the number of participants with each score in ARISTOTLE are summarised in Table 4.6. Potential solutions that would give a similar CHADS2 score and stroke risk factor distribution to ARISTOTLE could be defined as the number of patients with different combinations.

Simultaneous equations were derived relating the total number of patients having each characteristic to aid identification of potential solutions yielding the target distribution. Solutions were found via numerical optimisation and by applying any restrictions arising from the set of available patients in CPRD Aurum; more detail is given in the Appendix A2.3.2.

Table 4.6 $CHADS_2$ Score distribution in ARISTOTLE and associated risk factor combinations

$CHADS_2$ Score	Stroke risk factor(s) in age < 75 years	Stroke risk factor(s) in age ≥ 75 years
0 (N=54)	reduced LVEF and /or history of SE	N/A
1 (N=3046)	C or H or D	A
2 (N=3262)	CH or CD or HD or S	CA or HA or AD
3 (N=1681)	CHD or CS or HS or DS	CHA or CAD or HAD or AS
4 (N=767)	CHS or CDS or HDS	CHAD or CAS or HAS or ADS
5 (N=273)	CHDS	CHAS or CADS or HADS
6 (N=37)	N/A	CHADS

LVEF = left ventricle ejection fraction; SE = systemic embolism; N/A = not applicable;
C = congestive heart failure; H = hypertension; A = age 75 or older; D = diabetes; S = history of stroke or TIA.
Combinations of letters represents combinations of risk factors for example HS represents a person with hypertension AND prior stroke.
Target numbers (N=XX) derived from tabulations of baseline characteristics of ARISTOTLE participants.

Additional important variables – namely sex, prior VKA exposure status, and a more refined breakdown of age were added to the target multivariate covariate distribution. The distribution of stroke risk factors and age in the ARISTOTLE trial participants differed between men and women (133) with women older than men on average (median age 72 vs. 69) and a higher proportion of women having $CHADS_2$ score ≥3 (34.1% vs. 28.1% for men). This information was used to create separate target covariate distributions for women and men.

The idea of coarsened exact matching (CEM) was adapted wherein random sampling was used to select the required number in each subgroup from the CPRD Aurum patients available that would result in the target multivariate covariate distribution. The use of random sampling within subgroups based on combinations of important covariates or risk factors is used in the ('coarsened exact matching') method proposed by Iacus, King, and Porro (16) as an alternative to propensity score matching.

This method was successful in selecting a cohort of CPRD Aurum apixaban patients matched to the ARISTOTLE participants.

4.3.5. Additional information on step 3: matching warfarin users to apixaban users

4.3.5.1. Inclusion of prevalent warfarin users

ARISTOTLE stratified randomisation based on prior vitamin k antagonist (VKA) exposure with 57% of the participants having prior exposure to VKA, this leads a researcher to an important question: *To replicate ARISTOTLE should the proportion of new and prior users of VKA be matched to the trial in our cohort?*

In the literature review in Chapter 3 it was observed that the majority of studies in this therapeutic area comparing apixaban to warfarin included only new users of the treatments under study or required a washout window prior to the index date. Restricting to new users alone makes the selection of the cohort far simpler in that there is an obvious choice for the index date (the date of the first prescription for the treatment of interest). In the context of a reference trial emulation, where we also need to apply the trial eligibility criteria at the index date, the selection of index date has the potential to introduce selection bias (135). The bias arises firstly from the fact that prevalent users have 'survived' the initial post-initiation time period in which the risks of a drug may be higher, and secondly a risk that the covariates for the prevalent users (which are used to balance cohorts) may have been affected by the prior exposure to the drug of interest (135).

A simple solution for the emulation of ARISTOTLE would therefore have been to restrict the cohort to new users. The important question is then whether the treatment effect estimated in a cohort of new users alone is comparable to the treatment effect estimated in the trial, in which the majority of participants were prevalent users of VKA.

When we consider the choice in whether to match the trial on the inclusion of prevalent users we are confronted with thinking about the trial estimand and a more general question in

reference trial emulation – *what aspects of the trial are important/relevant to match and how comparable are the treatment estimates obtained depending on the choices made?*

For ARISTOTLE, prior VKA exposure was considered an important variable to match to ensure the treatment estimates obtained would be answering the same question. An interpretation of the different questions being answered by a study design including both new and prevalent users (Q1) vs the question answered by a study restricted to new users (Q2) is given below.

Q1: What is the treatment effect of initiating or switching to apixaban vs initiating or continuing on warfarin on time to stroke or systemic embolism in NVAF patients with 1 stroke risk factor?

This is a conceptualisation of the question being investigated in ARISTOTLE; this was judged as being materially different from the question that would be answered restricting to new users alone (Q2).

Q2: What is the treatment effect of initiating apixaban vs initiating warfarin on time to stroke or systemic embolism in NVAF patients with 1 stroke risk factor?

Although the difference may appear subtle, the two questions are measuring different estimands and one has relevance to both new and prevalent users (Q1) whereas the other does not (Q2). The question of whether to continue on a treatment or switch to an alternative is different from the question of which treatment to choose when treatment-naive.

Restricting to new users alone can result in a cohort not representative of 'real world' patients - when a new drug is introduced to the market as an alternative to an existing well-established treatment, the 'real world' cohort is most likely to consist of a large proportion of existing users of the older treatment switching to the new treatment. New user design studies which

exclude these patients are therefore limited in their generalisability and necessarily exclude a large proportion of patients.

Within the framework of this thesis in which the primary focus was the emulation of the reference trial and an exploration of the limits on the ability to emulate a trial, it was considered necessary to attempt to replicate this aspect of the trial design. ARISTOTLE randomisation was stratified on prior VKA exposure thus the matching of apixaban and warfarin users to be stratified in a similar way.

Having decided to include prevalent users the challenge was then determining how to select the index date for the prevalent warfarin users, how to apply the trial eligibility criteria to the prevalent warfarin users without introducing selection bias, and how to match on prior exposure to VKA and characterise this prior exposure.

A naïve approach to select the index date for the prevalent warfarin users would be to select either the first date or a random date in the treatment period for each prevalent warfarin user. These approaches introduce selection bias leading to biased treatment effect estimates (136). This is a particular problem in the emulation of trials where we may be assessing multiple eligibility criteria across a set of potential index dates for a patient.

4.3.5.2. Prevalent new user design

An important paper on the topic of including prevalent users is the Suissa 2017 (137) paper. This pivotal paper provided a framework for how to conduct pharmacoepidemiological studies including both new and prevalent users, a design Suissa named the ‘Prevalent New User’ (PNU) design. The method Suissa proposed involves constructing an ‘exposure set’ for each patient that switches to the newer study drug of interest, comprising all patients continuing on the comparator treatment that have the same history of prior treatment. Time-

conditional propensity scores are calculated by setting all exposure sets together into 1 dataset and using conditional logistic regression to estimate the probability of switching vs continuing within each exposure set. Patients are then selected by propensity score matching in chronological order (from earliest index date of a switcher to latest), with any patients selected as a match no longer eligible to be a match and dropped from subsequent exposure sets. More detail on this method is given in the Appendix.

This method appeared promising for the application of trial emulation and was applied to the ARISTOTLE emulation. In Suissa's PNU design the exposure sets are created with replacement therefore, applied to the ARISTOTLE emulation, a continuing warfarin user may be included in numerous, or even all possible exposure sets. A key challenge with the PNU design is the large size of the datasets involved; setting all exposure sets together gave a dataset of size approximately 10 billion records when applied to the CPRD Aurum patients prescribed apixaban and warfarin. An additional complication is the checking of eligibility criteria at the point of selecting a match (ie after calculation of the propensity scores) requiring the addition of variables relating to the eligibility criteria to the propensity score models. The conditional logistic regression to estimate the time-conditional propensity score models failed to converge thus an alternative approach to inclusion of prevalent users was sought.

4.3.5.3. Approach taken for including prevalent users

At the time of trialling the PNU method Webster Clarke et al published a paper (136) introducing alternative approaches inspired by the PNU approach but with the potential for easier implementation. One of these methods, 'forward random sampling of continuers', appeared particularly suitable; furthermore, this method demonstrated good results in their simulation study with the results being close to the true results (0.997) and performing nearly as well as the full PNU design of Suissa.

The ‘forward random sampling of continuers’ method applied to our study involved ordering the VKA-experienced apixaban users from shortest prior VKA treatment duration to longest prior VKA treatment duration, then for each VKA-experienced apixaban user in turn selecting a sample of 5 or so warfarin continuers with an equivalent duration of prior VKA treatment. The examples given by Webster-Clarke did not cover how to apply eligibility criteria; if we consider the RCT processes we are trying to emulate we have:

- Screening into an RCT at which point a potential participant may pass or fail
- Randomisation of a participant to treatment

When we apply the trial eligibility criteria to the CPRD Aurum patients this is analogous to the process of screening. Crucially we must not forget that in an RCT a participant can fail screening. A clinician may consider a patient as being potentially eligible if they meet the minimum inclusion criteria (in our case a patient aged minimum 18 years with AF diagnosis and at least one stroke risk factor), however they may be revealed as failing exclusion criteria during the screening process. Many of the exclusion criteria in ARISTOTLE are time dependent – such as recent abnormal lab values, recent stroke, use of certain concomitant medications, whereas others are binary with a participant becoming permanently ineligible after certain diagnoses or after certain surgeries such as placement of a mechanical heart valve. To most closely mimic the screening process for prevalent users we must therefore use a method in which it is possible for the patient to be selected at a date at which they might fail screening. This can be conceptualised as demanding that

i) our selection method should be ‘blind’ to the eligibility status of the prevalent user until after their index date has been selected
ii) should a patient ‘fail’ screening by having been selected at an index date at which they were not trial-eligible then they should not be able to be reselected at a different date; in other words each prevalent user should only have 1 chance at being selected in the same way that the apixaban users have only 1 chance to pass screening

One way to implement this requirement within the 'forward random sampling of continuers' method was to hide the eligibility status of the participants before sampling and check the status at the point of sampling. Should a patient be sampled at a date where they happen to be ineligible then that patient could then be considered as having failed screening and be removed from the pool of potential prevalent user matches.

The sampling algorithm was run first attempting to sample 5 continuing warfarin users per switcher to apixaban resulting in a suitable sample size available for each switcher; however there remained a large pool of warfarin continuers not sampled. The algorithm was therefore re-run sampling 10 continuers per switcher which successfully sampled continuing warfarin users even at the longer treatment durations.

After sampling, prevalent users were propensity score matched within treatment history 'strata' with a requirement for matches to have equivalent duration of prior exposure. Different options were trialled for the treatment history strata:

i) 3-strata [0-10 months, 10months-3 years, >3 years]
ii) 4-strata [0-4month, 4-12 months, 1-5 years, >5 years]
iii) 6-strata [0-4 months, 4-12 months, 1-2.5 years, 2.5-5 years, 5-10 years, >10 years]

The option that resulted in the largest and most balanced prevalent-user cohort was selected assessing balance by comparison of standardised mean differences of the baseline characteristics.

4.3.6. Additional information on sensitivity analyses

Initially, it was planned for cohort inclusion to require ≥2 prescriptions for the OAC of interest. This requirement was removed from the primary analysis over concerns this may introduce selection bias or immortal time bias(138). The analysis requiring ≥2 prescriptions was retained as an additional sensitivity analysis to understand the impact of including patients that may not have taken the index medication.

There was a high rate of missing data for INR control prior to the index date meaning this was not included in the propensity score models for the primary analysis. A sensitivity analysis was therefore performed including a variable on prior INR control.

4.4. Summary

This chapter presented the protocol for the emulation of ARISTOTLE using UK electronic healthcare records (CPRD Aurum linked to HES and ONS). The protocol paper was followed by a summary of the code lists used in the application of the trial eligibility criteria and derivation of TTR. This chapter presented the development of the method used in matching the baseline characteristics of ARISTOTLE using only publicly available aggregate summaries. A brief overview of 'prevalent new user design' methods was presented, followed by a description of the sampling method selected ('forward random sampling of continuers') and the adaptation made to emulate the screening process of an RCT.

The next chapter will present the results of the emulation of ARISTOTLE in CPRD Aurum including any deviations from the protocol-planned analyses and any post-hoc additional analyses performed.

Chapter 5 Research paper: Results

This chapter will present the results of the emulation of ARISTOTLE using data from CPRD Aurum linked to ONS and HES. The key results were presented in a results paper which is included here in section 5.1. After the paper additional results are presented in section 5.2. The results cover:

- The application of the eligibility criteria to the cohort, matching the apixaban users to ARISTOTLE on baseline characteristics, and propensity score matching to warfarin users
- A comparison of the results in the EHR cohort against the benchmarking criteria and reference trial results
- A post-hoc analysis looking at the quality of apixaban dose adjustment in the apixaban group in the ARISTOTLE-analogous EHR cohort
- Sensitivity analyses looking at the potential impact of treatment switching, moving the study start date a year later, and the application of a minimum exposure requirement.
- Results by prior VKA exposure strata
- Results of analysis by TTR and sensitivity analysis
- Results from an analysis of ARISTOTLE-eligible new users without matching to the ARISTOTLE baseline characteristics

5.1. Research paper: Results

5.1.1. Introduction to Paper 2

Summary

Chapter 4 described the methods used for the objective of emulating ARISTOTLE using CPRD Aurum linked to HES and ONS data. In this chapter I present the results paper for the emulation of ARISTOTLE using CPRD Aurum data. The results paper summarises the methods used and results of the study, including the primary effectiveness results, safety results, and sensitivity analyses. The results paper was under final review in May 2024 for publication in *PLOS Medicine*.

Thesis objectives addressed

This chapter describes the analyses that were planned in addressing the following objectives of the overall thesis (Section 1.3):

1. Emulate the reference trial ARISTOTLE comparing apixaban to warfarin for prevention of stroke in atrial fibrillation in UK EHRs including application of the trial eligibility criteria, matching to the baseline characteristics of the participants in the reference trial, and assessing the validity of the results and methods by benchmarking.

Role of candidate

I drafted the paper providing a brief summary of the methods used, the results obtained, and discussion and limitations. I performed the analyses including cohort selection and outcomes in the trial-analogous cohort. Kevin Wing (KW), Ian Douglas (ID), and Usha Gungabissoon (UG) provided guidance on appropriate handling of the data sources, methods, the discussion, and the strengths and limitations of the study. Liam Smeeth reviewed the manuscript. John Tazare (JT) provided guidance on prevalent new user (PNU) methods that had the potential to be suitable to the objective of this study and review of the manuscript. Turki Bin Hammad

(TBH) provided review of the ICD-10 codelists used as the outcomes and review of the manuscript. Angel Wong (AW) and Paris Baptiste (PB) provided review of the manuscript. The paper was finalised after review and suggested updates and comments from ID, UG, LS, JT, TBH, PB, AW, and KW. The paper was further updated following peer review to add more detail to the methods in the manuscript and limitations and discussion sections, to make an improvement to the analysis by TTR, and to add an analysis of the quality of apixaban dose-adjustment in the CPRD Aurum cohort.

5.1.2. Paper 2 coversheet

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1703768	**Title**	Ms
First Name(s)	Maud Emma Louise		
Surname/Family Name	Teoh		
Thesis Title	Real-world effectiveness of oral anticoagulants in the prevention of stroke: emulation and extension of the ARISTOTLE trial using UK EHRs		
Primary Supervisor	Kevin Wing		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	PLOS Medicine
Please list the paper's authors in the intended authorship order:	Emma Maud Powell, Usha Gungabissoon, John Tazare, Liam Smeeth, Paris J Baptiste, Turki M Bin Hammad, Angel YS Wong, Ian J Douglas, Kevin Wing

Stage of publication	**Undergoing revision**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I drafted the paper providing a brief summary of the methods used, the results obtained, and discussion and limitations. I performed the analyses including cohort selection and outcomes in the trial-analogous cohort. Kevin Wing (KW), Ian Douglas (ID), and Usha Gungabissoon (UG) provided guidance on appropriate handling of the data sources, methods, the discussion, and the strengths and limitations of the study. Liam Smeeth reviewed the manuscript. John Tazare (JT) provided guidance on prevalent new user (PNU) methods that had the potential to be suitable to the objective of this study and review of the manuscript. Turki provided review of the ICD-10 codelists used as the outcomes and review of the manuscript. Angel Wong and Paris Baptiste provided review of the manuscript. The paper was finalised after review and suggested updates and comments from ID, UG, LS, JT, TBH, PB, AW, and KW. The paper was further updated following peer review to add more detail to the methods in the manuscript and limitations and discussion sections, to make an improvement to the analysis by TTR, and to add an analysis of the quality of apixaban dose-adjustment in the CPRD Aurum cohort..

SECTION E

Student Signature	[redacted]
Date	12 April 2024

Supervisor Signature	[redacted]
Date	12 April 2024

5.1.3. Paper 2

Comparison of oral anticoagulants for stroke prevention in atrial fibrillation using the UK Clinical Practice Research Datalink Aurum: A reference trial (ARISTOTLE) emulation study

Emma Maud Powell[1]*, Usha Gungabissoon[2], John Tazare[3], Liam Smeeth[1], Paris J Baptiste[4], Turki M Bin Hammad[1,5], Angel YS Wong[1], Ian J Douglas[1], Kevin Wing[1]

[1] Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

[2] Epidemiology, GSK, London, United Kingdom

[3] Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

[4] Clinical Effectiveness Group, Centre for Primary Care, Wolfson Institute of Population Health, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom

[5] Methodology and biostatistics team, Department of Efficacy and Safety, Drug sector, Saudi Food and Drug Authority, Riyadh, Saudi Arabia

* Corresponding author

E-mail maud.teoh@lshtm.ac.uk (EMP)

ABSTRACT

Background

Stroke prevention treatment guidance for patients with atrial fibrillation (AF) uses evidence generated from randomised controlled trials (RCTs). However, applicability to patient groups excluded from trials remains unknown. Real-world patient data provides an opportunity to evaluate outcomes in a trial analogous population of direct oral anticoagulants (DOACs) users and in patients otherwise excluded from RCTs, however there remains uncertainty on the validity of the methods and suitability of the data.

Successful reference trial emulation can support the generation of evidence around treatment effects in groups excluded or underrepresented in the original trials.

We used linked UK primary care data to investigate whether we could emulate the pivotal ARISTOTLE trial (apixaban vs warfarin) and extend the analysis to investigate the impact of warfarin time in therapeutic range (TTR) on results.

Methods and findings

Patients with AF in a UK primary care database Clinical Practice Research Datalink (CPRD Aurum) prescribed apixaban or warfarin from 1 Jan 2013 to 31 Jul 2019 were selected. ARISTOTLE eligibility criteria were applied to this population and matched to the RCT apixaban arm on baseline characteristics creating a trial-analogous apixaban cohort; this was propensity-score matched to warfarin users in the CPRD Aurum. ARISTOTLE outcomes were assessed using Cox proportional hazards regression stratified by prior warfarin exposure status during 2.5 years of patient follow-up and results benchmarked against the trial results before treatment effectiveness was further evaluated based on (warfarin) time in therapeutic range (TTR).

The analysis sample comprised 8734 apixaban users and propensity-score matched 8734 warfarin users in CPRD. Results [Hazard Ratio (95% Confidence Interval)] confirmed apixaban non-inferiority for stroke or systemic embolism (SE) [CPRD 0.98 (0.82,1.19) vs trial 0.79 (0.66,0.95)] and death from any cause [CPRD 1.03 (0.93,1.14) vs trial 0.89 (0.80,0.998)] but did not indicate apixaban superiority. Absolute event rates for Stroke/SE were similar for apixaban in CPRD Aurum and ARISTOTLE (1.27%/year) whereas a lower event rate was observed for warfarin (CPRD Aurum 1.29%/year, ARISTOTLE 1.60%/year)

Analysis by TTR suggested similar effectiveness of apixaban compared with poorly controlled warfarin (TTR < 0.75) for Stroke/SE [0.91 (0.73,1.14)], all-cause death [0.94 (0.84,1.06)], and superiority for major bleeding [0.74 (0.63, 0.86)]. However, when compared with well controlled warfarin (TTR ≥ 0.75) apixaban was associated with an increased hazard for all-cause death [1.20 (1.04, 1.37)] and there was no significant benefit for major bleeding [1.08 (0.90-1.30)]. The main limitation of the study's methodology are the risk of residual confounding, channelling bias and attrition bias in the warfarin arm, and selection bias and misclassification in the analysis by TTR.

Conclusions

Analysis of non-interventional data generated results demonstrating non-inferiority of apixaban vs warfarin consistent with the pre-specified benchmarking criteria. Unlike in ARISTOTLE superiority of apixaban vs warfarin was not seen which may be linked to the lower proportion of Asian patients and higher proportion of patients with well-controlled warfarin compared to ARISTOTLE. The methodological template developed can be used to investigate treatment effects of oral anticoagulants in patient groups excluded from or under-represented in trials and also provides a framework which can be adapted to investigate treatment effects for other conditions.

AUTHOR SUMMARY

Why Was This Study Done?

- Stroke prevention treatment guidelines for patients with atrial fibrillation (AF) are based on results from randomised controlled trials (RCTs), we do not know if these results are relevant to patients that would not have been eligible to be included in the RCTs.
- This study used routinely collected health data from the UK to emulate an RCT that compared apixaban to warfarin, ARISTOTLE, and also looked at whether the benefit of apixaban compared with warfarin was impacted by the quality of warfarin therapy (measured by time in therapeutic range, TTR).
- Emulating an RCT for stroke prevention in patients with AF should help to understand how transferable RCT results are to 'real-world' practices and whether this methodological approach can help to improve treatment options and outcomes for patient groups currently underrepresented in clinical trials.

What Did the Researchers Do and Find?

- The researchers looked at patients with AF in a UK primary care data prescribed apixaban or warfarin and applied a "reference trial emulation" approach, in which the ARISTOTLE trial eligibility, selection and analysis approaches were applied to UK primary care data and results benchmarked against those of ARISTOTLE.
- Patients prescribed apixaban had similar rates of outcomes to those prescribed warfarin in our cohort and our results were successfully benchmarked against ARISTOTLE. Unlike ARISTOTLE we did not see superiority of apixaban vs

warfarin [Hazard ratio (95% confidence interval)] for time to stroke or systemic embolism: 0.98 (0.82,1.19) in our cohort vs 0.79 (0.66,0.95) in ARISTOTLE.

- We also found the benefit of apixaban vs warfarin differed for some outcomes depending on the quality of warfarin therapy with apixaban (i) superior only to poorly controlled warfarin therapy for major bleeding [TTR <0.75: 0.74 (0.63, 0.86), TTR ≥ 0.75: 1.08 (0.90, 1.30)] (ii) associated with an increased risk of death compared only to well-controlled warfarin therapy [TTR ≥ 0.75: 1.20 (1.04, 1.37), TTR < 0.75: 0.94 (0.84, 1.06)].

What Do These Findings Mean?

- Our results support the NICE guidelines on selecting treatment for stroke prevention in patients with AF and also provide reassurance on continuing warfarin in patients with high TTR.
- We can use UK primary health care data to emulate a reference trial of treatments for the prevention of stroke in AF.
- We can use the data and methods to look at how well treatments work in patients that would not have been included in RCTs such as those with multimorbidity or patient groups under-represented in RCTs such as ethnic minority groups and older patients.
- Study limitations include the possibility of residual confounding, a risk that patients doing well on warfarin were over-represented in our cohort, a lower proportion of Asian participants in our cohort compared with ARISTOTLE, and the likelihood of residual selection bias/misclassification in the TTR analysis.

Introduction

Atrial fibrillation (AF) is a common type of cardiac arrhythmia with an estimated prevalence of 3.3% in UK adults aged ≥ 35 years [1]. AF is a risk factor for stroke; patients with AF have a five-fold increased risk of stroke compared with people without AF [2] and around a quarter of all strokes are attributed to this arrhythmia [3]. In addition, increased levels of mortality, morbidity and disability with longer hospital stays are observed in stroke patients with AF compared with stroke patients without AF [4, 5].

Pharmacological therapy recommended to reduce the risk of stroke in AF includes the use of oral anticoagulants (OACs). The introduction of direct oral anticoagulants (DOACs) for AF since 2012 in the UK provided a choice of treatment alongside the older OAC class of vitamin K antagonists (VKA), such as warfarin which has been available for over 60 years. The VKA OACs require regular monitoring of international normalised ratio (INR) to keep patients in the optimal therapeutic range (typically 2.0 to 3.0) in which risk of both ischemic and bleeding events are minimised [6]. A patient may require dose adjustments to stay within their INR target range. A key measure of quality of warfarin treatment is therefore the time in therapeutic range (TTR) which estimates the proportion of time a patient has spent with INR within optimal range. A TTR of 0.75 or greater is often considered as indicating optimal INR control and suggests a patient is spending a high proportion of their time in their INR target range.

ARISTOTLE was a pivotal RCT of the DOAC apixaban designed to demonstrate non-inferiority compared with warfarin in the prevention of stroke or systemic embolism (SE) in patients with AF. The results demonstrated superiority of apixaban over warfarin for both prevention of stroke/SE and safety (major bleeding) [7]. Results in the EU patient subset from the trial suggested the observed superiority of apixaban might be dependent on how

well warfarin therapy was managed in the comparator group [8], an analysis that has not yet been performed outside of trial settings. In the NICE review of ARISTOTLE, several professional groups noted the TTR of warfarin users in ARISTOTLE may be lower than what is typical in UK clinical practice [9].

Treatment guidelines for DOACs are based on evidence from randomised controlled trials (RCTs), however, it is unclear whether these results extend to patient groups typically excluded from trials such as those with increased bleeding risk or severe comorbidities. Whilst there have been a number of previous studies of DOAC effectiveness using non-interventional data, there remains uncertainty on whether the data sources and methods used have fully accounted for the lack of treatment randomisation and issues such as selection bias and confounding. Comparing results from real-world studies with RCT results is challenging due to differences in patient populations, treatment adherence, and study design. However, reference trial emulation involves use of an existing named RCT to (1) inform observational study design and (2) benchmark results against, providing confidence in validity of the selected observational methods and data. [10-13]. The non-interventional analysis methods can then be applied, under a set of assumptions, to reliably estimate effects in groups of patients with AF who would have been excluded from (or underrepresented in) the reference trial [14] such as patients aged > 80 that were under-represented in ARISTOTLE compared with patients with AF in UK clinical practice and patients with increased bleeding risk that were excluded by the trial eligibility criteria.

There is increasing interest in trial emulation using observation data, and in the application of recent developments in pharmacoepidemiology methods involving the inclusion of prevalent users. This study used a framework which involved coarsened exact matching to select patients matching the trial population on aggregate, and sampling of prevalent users in a way that avoids selection bias and emulates the process of screening into an RCT, to construct a

cohort of patients similar to the target trial population which included both new and prevalent users. This methodological approach could be adapted to a variety of treatments and different therapeutic areas.

This study sought to (1) create an ARISTOTLE-analogous cohort using routinely collected primary and secondary care data in the UK and (2) benchmark results obtained in the ARISTOTLE-analogous cohort with ARISTOTLE results and (3) explore whether apixaban treatment-effects in clinical practice are influenced by how well warfarin therapy is controlled.

Materials and methods

This study is reported as per the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline (S1 Checklist).

Study design

A propensity score matched cohort study with emulation of a reference trial (ARISTOTLE).

Setting/data sources

UK Electronic Healthcare Records

This study used non-interventional data from UK Clinical Practice Research Datalink (CPRD) Aurum, a database containing anonymised data from 738 primary care practices across England (approximately 13% of the population of England with 19 million patient records and 7 million active as of September 2018 [15]. CPRD Aurum contains information on clinical diagnoses, prescribing, referrals, tests, and demographic/lifestyle factors and is representative of the population of England in geographical spread, social deprivation, age, and sex [15]. This study also used 2 additional data sources linked to CPRD Aurum: Hospital

Episodes Statistics (HES) data, which contains data on patients admitted to NHS hospitals including diagnoses, admission and discharge, and Office of National Statistics (ONS) mortality data.

The reference trial (ARISTOTLE)

ARISTOTLE was a randomised, double-blind trial completed in 2011, comparing apixaban with warfarin in the prevention of stroke and SE. The trial included 18201 patients with AF and at least one additional risk factor for stroke. The trial was designed to test for non-inferiority of apixaban compared with warfarin (non-inferiority margin of 1.38 for the upper limit of the 95% CI of the hazard ratio for the primary outcome), and showed apixaban superiority for (1) the primary outcome of stroke or SE (HR 0.79; 95% CI 0.66, 0.95),7 (2) the safety endpoint of major bleeding (HR 0.69; 95% CI 0.60, 0.80), and (3) death from any cause (HR 0.89; 95% CI 0.80, 0.99). The ARISTOTLE findings led to the National Institute for Health and Care Excellence (NICE) guidelines on stroke prophylaxis in patients with AF recommending apixaban as a treatment.

ARISTOTLE eligibility criteria and summary baseline patient characteristics were used to select a cohort of patients from CPRD Aurum analogous to the ARISTOTLE participants.

The use of CPRD and ARISTOTLE are described in a previous publication [14] and use of CPRD for this project was approved by the MHRA Independent Scientific Advisory Committee [ISAC protocol in S2]. All data used in this study were anonymised.

Diagnostic and therapeutic codelists

All diagnostic and therapeutic codelist files used are available at https://datacompass.lshtm.ac.uk/id/eprint/3590/.

Patient Selection

Step 1: application of trial eligibility criteria to patients in CPRD

We first selected HES-linked patients registered in CPRD Aurum between January 1, 2013 and July 31, 2019, who had at least 6 months between registration and the index date. ARISTOTLE recruited both new (warfarin-naïve) and prevalent (warfarin-experienced) users of warfarin with randomisation stratified on prior warfarin (or other VKA) exposure status (warfarin naïve or experienced). To be classified as warfarin-naïve patients were required to have no evidence of exposure to warfarin or other VKA in the 5 years prior to the index date. To enable selection of a similar cohort of patients in CPRD Aurum (including both new and prevalent users of warfarin), the following process was used in determining index date:

- <u>apixaban users</u>

  index date = first prescription of apixaban in the study period

  apixaban user classified as warfarin-naïve or warfarin-experienced at this date;

- <u>warfarin users</u>

  for new users of warfarin: index date = first prescription of warfarin in the study period;

  for prevalent users of warfarin: a pool of potential index dates was selected containing all prescription dates in the study period, with index date selected at the later treatment-history sampling stage (see step 3).

ARISTOTLE eligibility criteria (supplementary table A2) [7] were applied giving a trial-eligible cohort for apixaban users, a trial-eligible cohort of new users of warfarin, and a pool of potential index dates (with all potential index dates kept in regardless of ARISTOTLE eligibility at this stage) for warfarin continuers (prevalent warfarin users).

Step 2: selection of apixaban trial-analogous patients in CPRD

We selected a subset of the CPRD Aurum trial-eligible apixaban patients that better matched the ARISTOTLE apixaban participants based on aggregate summaries for the following key ARISTOTLE baseline characteristics:

- Age
- Sex
- Congestive heart failure or left ventricular systolic dysfunction
- Hypertension requiring treatment
- Diabetes mellitus
- Prior stroke/transient ischemic attack (TIA)/systemic embolism (SE)
- Level of renal impairment
- Prior VKA/warfarin exposure

To characterise the baseline patient characteristics of ARISTOTLE, we used the key publication of the trial results [7], discussion of trial results by regulatory bodies [8, 9, 16] and publications on the trial presenting cross-tabulations on key characteristics [17,18].

An ARISTOTLE-analogous cohort of CPRD Aurum apixaban patients was then selected using a modified form of coarsened exact matching [19] (see Appendix for details).

Step 3: matching of apixaban trial-analogous patients to warfarin trial-eligible patients in CPRD

To emulate ARISTOTLE which stratified randomisation on prior VKA exposure status, patients in the CPRD cohort were matched separately within the VKA-naïve and VKA-experienced strata. A 3-step procedure, based on methods proposed by Suissa et al [20] and Webster Clark et al [21], was used to select and match patients in the VKA-experienced strata

whilst avoiding selection bias; this procedure is summarised in Figure 1 and described in in S3 Appendix.

The trial-analogous CPRD Aurum apixaban patients were matched to warfarin CPRD Aurum patients using greedy nearest-neighbour matching on the logit of the propensity score (PS); a caliper of 0.2 times the standard deviation of the logit of the propensity score was used for matching as recommended by Austin [22].

The covariates included in the propensity score models are detailed Table 1.

Table 1: Covariates Included in the Propensity Score Models

Category	Variable List
Demographics	age, sex, ethnicity
$CHADS_2$ stroke risk factors	congestive heart failure or left ventricular systolic dysfunction, hypertension requiring treatment, diabetes mellitus, prior stroke/TIA/systemic embolism
Vascular stroke risk factors	prior myocardial infarction, peripheral artery disease, aortic plaque, history of pulmonary embolism or deep vein thrombosis
Other risk factors	body mass index, systolic blood pressure, history of bleeding, smoking status, alcohol consumption, socioeconomic status (imd2105_5), ethnicity
Concomitant medications	aspirin, clopidogrel, NSAIDs, antacids, statins, angiotensin-converting enzyme inhibitors (ACEIs) or angiotensin receptor blockers (ARBs), beta blockers, calcium channel blockers, statins, amiodarone, digoxin, proton pump inhibitors, H2 receptor antagonist
Comorbidities	renal function, history of fall, Charlson comorbidity components [COPD, connective tissue disease, peptic ulcer disease, liver disease, hemiplegia, cancer, haematological cancer], healthcare utilization [number of GP consults in the prior year, number of hospitalizations in the prior year]
AF factors	time since AF diagnosis, history of valvular disease, history of valvular surgery
Healthcare utilisation	number of GP consults in the prior year, number of hospitalizations in the prior year

AF=atrial fibrillation; COPD=chronic obstructive pulmonary disease; GP=general practitioner; NSAIDs=non-steroidal anti-inflammatory drugs; TIA=transient ischemic attack;

The model resulting in the most balanced cohort was chosen with balance assessed by looking at standardised differences across all variables after matching using a target threshold

of 0.05 for the maximum difference allowed for any individual variable. Balance of covariates considered to be most important in predicting outcome were prioritised namely age, sex, and stroke risk factors.

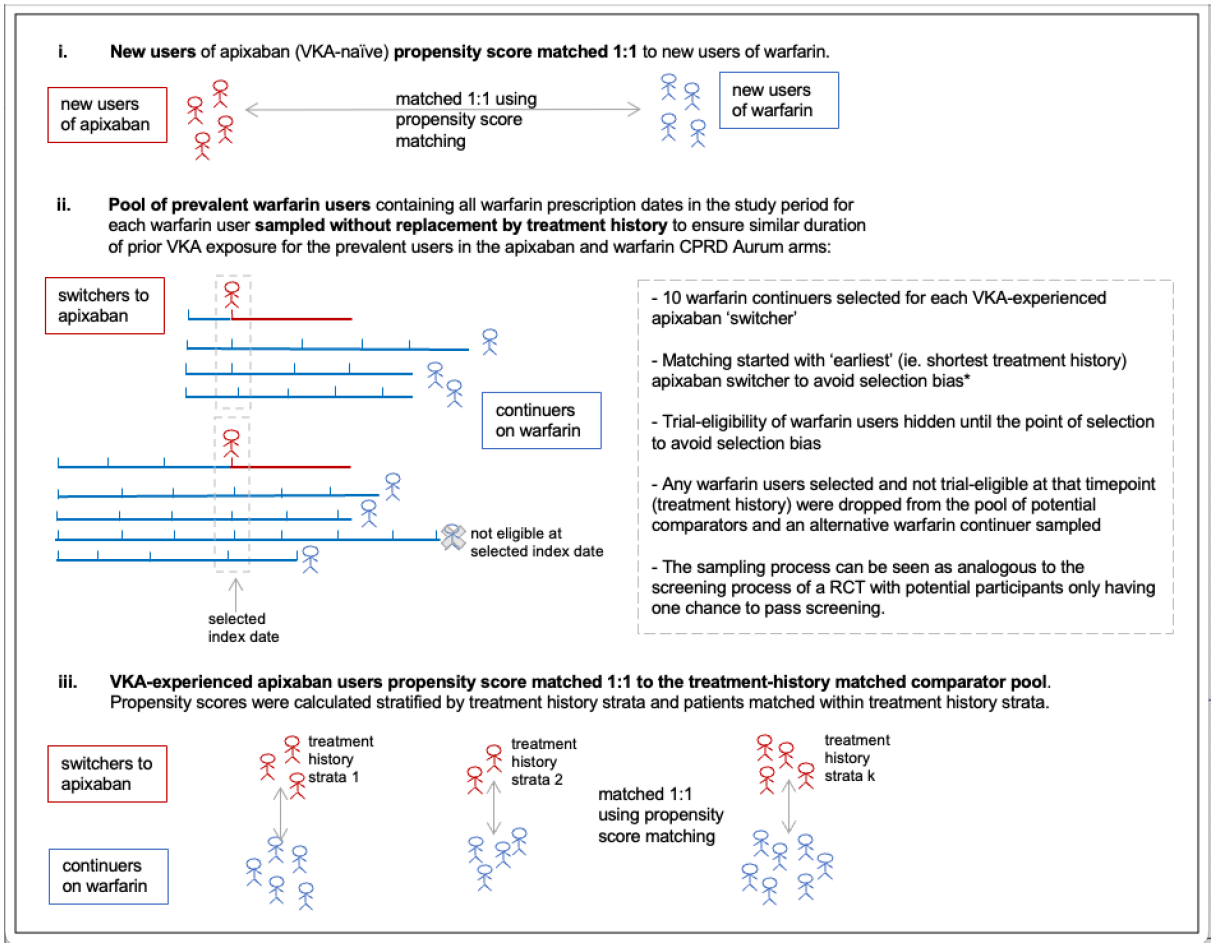


Figure 1: Matching of apixaban trial-analogous patients to warfarin trial-eligible patients
* This method has been found in a simulation study (Webster-Clark et al [21]) to give unbiased results
CPRD=Clinical Practice Research Datalink; RCT=randomised controlled trial; VKA=vitamin K antagonist.

Exposures and outcomes

Exposures

Exposure to apixaban (5mg/2.5mg) or warfarin was determined using CPRD prescribing records with no restrictions on the dose prescribed.

Outcomes

The primary effectiveness outcome was the composite of stroke (ischemic or haemorrhagic) or systemic embolism (SE); individual components of this outcome (stroke, ischemic or uncertain type of stroke, haemorrhagic stroke, SE) and death from any cause were the key secondary effectiveness outcomes. Secondary effectiveness outcomes included myocardial infarction (MI), pulmonary embolism or DVT, and composite endpoints of effectiveness outcomes. The primary safety outcome was major bleeding (including by location – intracranial, gastrointestinal, or other location such as urinary or gynaecological). All outcomes involved hospitalisation or death and were ascertained using HES and ONS data. The ICD-10 codes used in ascertaining stroke occurrence have been recommended as having high positive predictive value [23].

Statistical analysis

Methods of Analysis

A prospective protocol was published prior to the analysis detailing the planned analyses [14, also in Appendix].

Changes from the planned protocol are described in Table 2.

Table 2: Changes from Planned Analyses

Original Planned Analysis	Updated Analysis	Reason for change
Patients to be selected from both CPRD GOLD and CPRD Aurum.	Only CPRD Aurum used.	There was a much larger sample size available in CPRD Aurum meaning combining of the 2 data sources was not

		required.
Censoring scheme to censor at 5 years after index date.	Censoring scheme censored at 2.5 years after index date.	The ARISTOTLE trial had median duration of follow-up of 1.8 years (IQR 1.4, 2.3) therefore a 2.5 year cut-off gives a more similar duration of follow-up than 5 years.
Adherence of apixaban users to be measured by proportion of days covered by prescriptions.	Treatment persistence measured instead (proportion of patients still on index treatment at date of censoring).	Repeat prescriptions are often issued automatically meaning comparing number of days covered by prescribed pills to the number of days in the treatment period did not provide useful insight on adherence.
Supplementary analysis in patients deemed adherent (PDC ≥ 80%, ARISTOTLE compliance limit).	Analysis by TTR only.	Unable to ascertain useful measure of adherence in the apixaban users.
Non-inferiority will be concluded when the upper limit of the 95% CI for the hazard ratio must be less than 1.52 (upper limit in the EU subgroup of ARISTOTLE).	Non-inferiority will be concluded when the upper limit of the 95% CI for the hazard ratio is less than 1.38 (same non-inferiority margin of ARISTOTLE).	The non-inferiority margin used in ARISTOTLE was the one agreed by regulators to represent the maximum acceptable clinical difference. By applying the same margin, we ensure that the conclusion is based on more rigorous criteria.
Aim to include prior INR control in propensity model for vitamin k antagonist-experienced patients.	Primary analysis does not include prior INR control. Post hoc sensitivity analysis performed including prior INR control in the propensity score model.	High rate of missing data for prior INR control made it not advisable to include this variable in the propensity score model for the main analysis. Other variables predictive of poor INR control such as age are already included. Post hoc sensitivity analysis including INR control in the propensity score model performed to assess the potential impact of not including this variable following question in peer review on the omission of this variable.
N/A	Post hoc analysis assessing apixaban dose-adjustment in CPRD Aurum	Suggested by peer review to provide evidence on the quality of dose adjustment in CPRD Aurum and how this may impact the results in the trial-analogous cohort.

All time-to-event endpoints were analysed using a Cox proportional hazards model, stratified by prior VKA status (experienced, naïve). The effectiveness outcomes were analysed using the intention-to-treat principle and major bleeding was analysed using an on-treatment censoring scheme. Patients were censored at 2.5 years after index date reflecting typical maximum duration of follow-up in ARISTOTLE. Cluster-robust standard errors were used with pair membership as the clustering variable [24,25]. The proportional hazards assumption

was assessed by looking at the log-log of the Kaplan-Meier survival curves and inspection of scaled Schoenfeld residuals plotted against time. Analyses were performed using SAS version 9.4 and R version 4.2.1.

Supplementary analyses

A protocol planned analysis in the subset of patients deemed adherent (with adherence measured by TTR in the warfarin users and by proportion of days covered by prescriptions in the apixaban users) was planned to assess the impact of adherence on outcomes. The planned analysis was not possible due to the apixaban prescription data not providing a useful measure of adherence. . An analysis by INR TTR was performed instead to assess the impact of warfarin control on results with all outcomes analysed by TTR (TTR < 0.75 and TTR ≥ 0.75). Individual predicted TTR based on baseline variables was used for patients missing TTR. In order to perform the TTR analysis whilst maintaining balance in the baseline covariates, inverse probability treatment weighting (IPTW) was used to rebalance the baseline characteristics, applying stabilised weights to the ARISTOTLE-analogous apixaban users. A similar approach to the main analysis was used with propensity score models constructed separately for the new users and warfarin-experienced users.

An additional post hoc analysis was performed looking at the proportion of apixaban patients prescribed reduced-dose apixaban along with a comparison of the patients meeting the criteria for dose-reduction against the dose actually prescribed. In this analysis apixaban dose in the ARISTOTLE-analogous CPRD cohort was assessed and compared against the ARISTOTLE protocol-specified criteria and NICE criteria for reduced apixaban dose. ARISTOTLE specified that participants meeting any 2 of the following criteria assessed at the time of randomisation should have their apixaban dose reduced to 2.5mg BID: age ≥ 80 years, body weight ≤ 60 kg, or serum creatinine ≥ 1.5 mg/dL. These criteria are equivalent to

the NICE guidelines for dose reduction with NICE having an additional criteria indicating reduced dose in those with creatinine clearance 15–29 mL/minute.

In addition, to assess the impact of the quality of dose-adjustment in the CPRD cohort on the observed effectiveness of apixaban relative to warfarin, a supplementary post hoc analysis was performed looking at the results in the subset of apixaban patients prescribed the correct dose compared with IPTW re-balanced warfarin comparators.

Sensitivity analyses

Primary and secondary effectiveness outcomes were also analysed using the on-treatment censoring scheme to investigate whether treatment discontinuation compromises confidence in the effectiveness analyses.

Treatment persistence was defined by looking at longitudinal prescription data for OACs; OAC treatment windows were derived in which gaps >= 6 months between prescription dates were considered as distinct treatment windows. The end of each OAC treatment window was derived as the date of the last prescription of index OAC + the number of days supply given in the last prescription + a grace period of 30 days. In cases of overlapping OAC treatment windows the date of the first prescription of the subsequent OAC treatment window was used to define the end of the prior OAC window. A prescription for a different OAC from the index OAC treatment was considered as a treatment switch. An ending of index OAC treatment with no subsequent prescription for any other OAC recorded was considered as treatment stop. Gaps of >= 6 months with no subsequent OAC prescriptions recorded were categorised as having stopped OAC treatment.

The set of patients who switched or discontinued treatment during follow-up were examined to ascertain whether selection bias due to attrition may have affected the on-treatment analyses (Table A9 in Appendix).

Apixaban was first launched for AF in the UK in January 2013, with relatively few patients receiving a prescription in the first year it was available; we therefore performed a sensitivity analysis with the start of study period shifted forwards a year to investigate the impact of inclusion of early adopters who may differ from later adopters of a new drug.

Confounding and bias

In the study period apixaban was a newly available treatment leading to the possibility of channelling bias [26]. By applying trial eligibility criteria to both treatment cohorts and matching using baseline covariates we aimed to minimise channelling bias. To handle confounding, treatment arms were matched using PSM [27].

Benchmarking results against ARISTOTLE

The study hypothesis was that results in the CPRD ARISTOTLE-analogous cohort would be comparable to the ARISTOTLE results, as defined by the pre-specified benchmarking criteria. A slightly weaker benefit of apixaban vs warfarin was expected based on the weaker benefit seen in the EU subgroup of ARISTOTLE and an expectation that the quality of warfarin control in UK patients may be higher than that observed in ARISTOTLE.
The benchmarking criteria for considering the results in the trial-analogous CPRD cohort to be comparable with ARISTOTLE were pre-specified and published previously [14]:

1. The effect size must be clinically comparable with the ARISTOTLE findings; the HR for time to stroke/SE with the HR must be between 0.69 and 0.99. This range is not symmetrical around the ARISTOTLE estimate of 0.79 as it is anticipated the treatment effect in routine clinical care may be weaker than that seen in the optimised setting of a clinical trial.

2. The upper limit of the 95% CI for the HR for time to stroke/SE must be less than 1.38 (non-inferiority margin used in ARISTOTLE, updated since protocol – see Table 2).

The benchmarking step applied only to the primary effectiveness outcome in the trial-analogous CPRD cohort; results in other groups such as patients underrepresented or excluded from the trial would not necessarily be expected to remain consistent to the RCT results given the relative risks may differ in these groups. Comparability of other outcomes was to be assessed descriptively with no formal criteria or hypothesis testing used.

Missing data

Patients with missing systolic blood pressure (0.1%), body mass index (3.3%), smoking status (0.1%), or socioeconomic status (0.1%) were excluded from the trial-eligible cohort as the proportion of patients with these missing was low. Patients with missing renal function (1.3%), ethnicity (0.4%), or alcohol use (5.6%) were kept in the cohort through a missing indicator approach; this approach is valid under the assumption that these variables act as confounders and influence clinician prescribing decisions only when observed [28]. A total of 1176 (13.3%) warfarin users in the CPRD cohort did not have INR measurements in the data during their treatment period with predicted TTR used for these patients in the analysis by TTR (see Appendix for details).

Ethics

Scientific approval was provided by the London School of Hygiene and Tropical Medicine research ethics committee (ref 17682) and the independent scientific advisory committee of the Medicines and Healthcare Products Regulatory Agency (protocol no. 19_066R). CPRD data are already approved via a national research ethics committee for purely non-interventional research of this type. CPRD data are analysed anonymously therefore individual patient consent is not sought by contributing medical practices when data is shared

with CPRD; however, patients are able to opt out of their patient information being shared for research.

Results

Participants

Between January 1, 2013 and July 31, 2019 there were 86,888 people with AF prescribed apixaban and 159,632 prescribed warfarin in HES-linked CPRD Aurum practices (Figure 2). Application of minimum registration period and ARISTOTLE inclusion criteria reduced this to 67,539 apixaban and 139,527 warfarin patients. After applying ARISTOTLE exclusion criteria there were 41,487 apixaban and 101,159 warfarin patients.

Selecting apixaban patients to match ARISTOTLE on key baseline characteristics yielded 9,120 apixaban patients (3,912 new users and 5,208 prevalent users) available for propensity score matching to 101,159 warfarin patients. For 274 apixaban patients no match could be found giving a propensity score matched cohort of 8846 apixaban and 8846 warfarin patients.

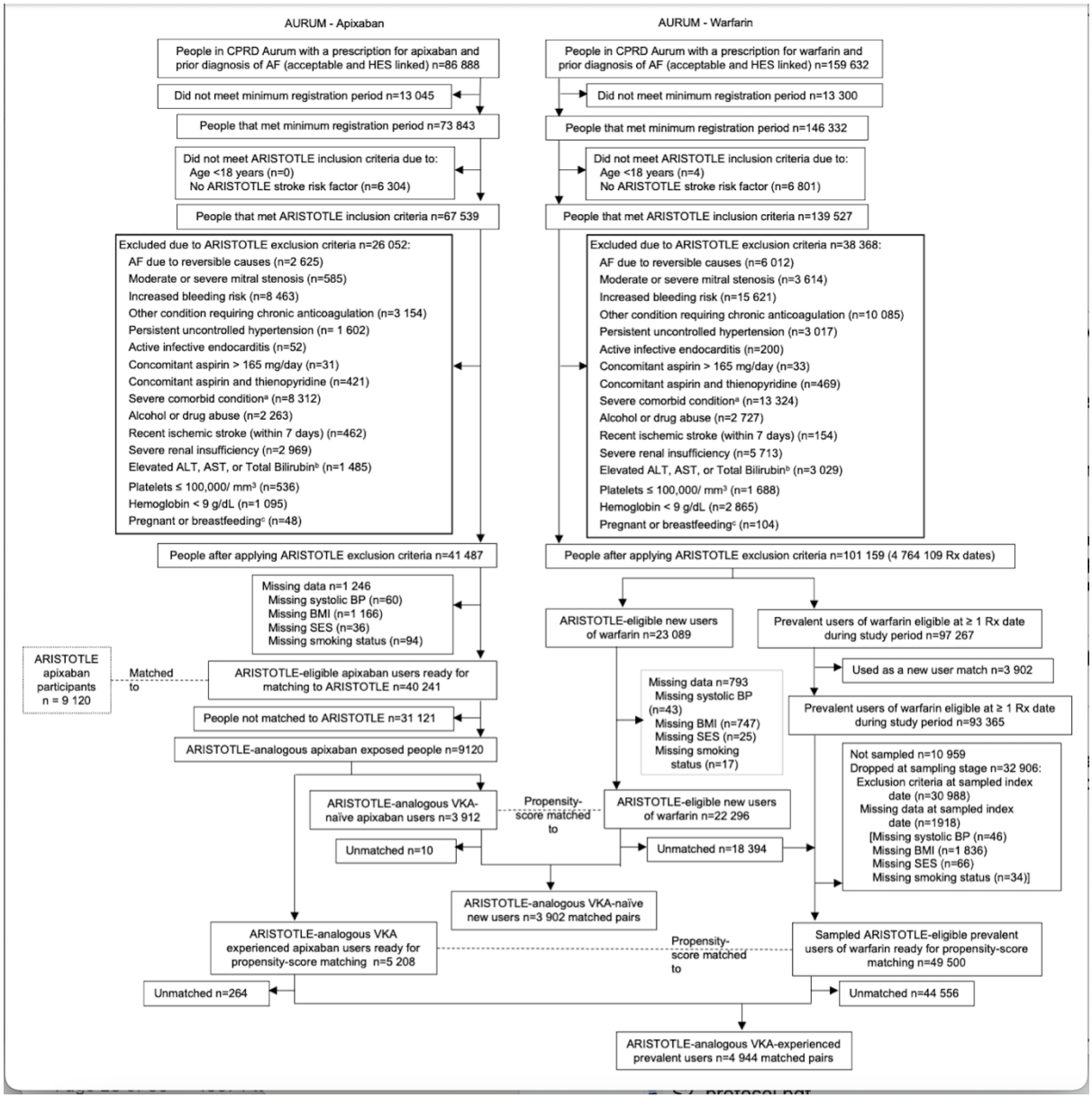


Figure 2: Selection of ARISTOTLE-analogous CPRD Aurum Cohort

Flow of number of individuals included in the analysis. AF = atrial fibrillation; ALT = alanine transaminase; AST = aspartate transaminase; BMI = body mass index; BP = blood pressure; CPRD = Clinical Practice Research Datalink; HES: Hospital Episodes Statistics; Rx = Prescription; SES = socioeconomic status; ULN = upper limit of normal; VKA = vitamin K antagonist.
a Severe comorbid condition with life expectancy <1 year or reasons making participation impractical; b ALT or AST > 2X ULN or Total Bilirubin ≥ 1.5X ULN; c Pregnant or breastfeeding within 3 years prior
See supplementary table A1 in S3 Appendix for detailed list of inclusion and exclusion criteria.
Note: For prevalent warfarin users trial eligibility only revealed at point of random selection into the cohort for prevalent users. Numbers in figure show maximum theoretical number of warfarin users available should they be selected only at a time they were eligible for the trial.

Application of ARISTOTLE inclusion/exclusion criteria and matching to ARISTOTLE

Applying the ARISTOTLE inclusion/exclusion criteria and matching to ARISTOTLE baseline patient characteristics resulted in a cohort similar to the ARISTOTLE apixaban participants (Table 3); for example median age was 78 and mean $CHADS_2$ score 2.4 in CPRD Aurum before applying trial criteria and matching whereas the median age of 71 and mean $CHADS_2$ score 2.1 after these steps matched the ARISTOTLE apixaban participants. The ARISTOTLE-analogous apixaban arm matched the trial arm on prior VKA exposure, age, sex, stroke risk factors and $CHADS_2$ score, and proportion of patients with moderate or severe renal impairment.

Differences remained on baseline characteristics it was not feasible to match on namely: ethnicity (95.2% white, 2.4% Asian in CPRD Aurum apixaban vs 82.6% white, 14.4% Asian in ARISTOTLE) and concomitant medications (amiodarone 3.8%, aspirin 5.8%, digoxin 13.9% in CPRD Aurum apixaban users vs amiodarone 11.1%, aspirin 31.3%, digoxin 32.0% in ARISTOTLE apixaban arm). See Appendix for details on matching feasibility.

Propensity score matching of CPRD Aurum trial-analogous apixaban users to CPRD Aurum warfarin users

Results of Propensity score matching

Before propensity score matching, differences between treatment groups were evident for most baseline variables including age (median age 71 in apixaban vs 78 in warfarin), sex (apixaban 35.6% female vs warfarin 43.6%), and stroke risk factors [see Table 3]. After propensity score matching all baseline characteristics were well balanced (maximum standardised difference 0.031). From 9120 apixaban users only 274 (3.0%) were dropped due to unsuccessful matching.

	CPRD Aurum							ARISTOTLE Trial	
	No ARISTOTLE criteria or matching		After applying ARISTOTLE criteria		After applying ARISTOTLE criteria, matching to the trial and PSM apixaban to warfarin				
Characteristic - n(%) unless specified	Apixaban (N=73 843)	Warfarin (N=146 332)	Apixaban (N=41 487)	Warfarin (N=101 159)	Apixaban (N=8 846)	Warfarin (N=8 846)	Standardised difference	Apixaban (N=9 120)	Warfarin (N=9 081)
Age – years , median (IQR)	78 (70, 85)	78 (71, 84)	78 (71, 84)	78 (72, 84)	71 (63, 77)	71 (63, 77)	0.008	70 (63, 76)	70 (63, 76)
Female sex	34 430 (46.6)	63 321 (43.3)	19 591 (47.2)	44 197 (43.7)	3144 (35.5)	3190 (36.1)	0.011	3 234 (35.5)	3 182 (35.0)
Systolic blood pressure – mmHg, median (IQR)	130 (120, 140)	130 (120, 140)	131 (120, 140)	130 (120, 140)	130 (120, 140)	130 (120, 140)	0.001	130 (120, 140)	130 (120, 140)
Missing	132	267	60	125	0	0			
Weight – kg, median (IQR)	79 (67, 92)	80 (68, 93)	80 (68, 93)	80 (69, 94)	85 (73, 100)	85 (74, 99)	0.003	82 (70, 96)	82 (70, 95)
Prior myocardial infarction	9 958 (13.5)	20 406 (13.9)	5 035 (12.1)	13 446 (13.3)	1090 (12.3)	1074 (12.1)	0.006	1319(14.5)	1266 (13.9)
Prior clinically relevant or spontaneous bleeding	16 972 (23.0)	31 034 (21.2)	7 721 (18.6)	19 007 (18.8)	1533 (17.3)	1507 (17.0)	0.008	1525 (16.7)	1515 (16.7)
History of fall within previous year	2 443 (3.3)	2 688 (1.8)	1 093 (2.6)	1 561 (1.5)	137 (1.5)	131 (1.5)	0.006	386 (4.2)	367 (4.0)
Prior use VKA >30 days	24 240 (32.8)	102 725 (70.2)	12 558 (30.3)	75 787 (74.9)	4944 (55.9)	4944 (55.9)	0.000	5 208 (57.1)	5 193 (57.2)
Qualifying risk factors									
Age ≥ 75 years	45 762 (62.0)	93 436 (63.9)	26 730 (64.4)	68 197 (67.4)	2 770 (31.3)	2 740 (31.0)	0.007	2 850 (31.2)	2 828 (31.1)
Prior stroke, TIA, or SE	20 713 (28.1)	38 132 (26.1)	11 422 (27.5)	25 898 (25.6)	1 711 (19.3)	1 709 (19.3)	0.001	1 748 (19.2)	1 790 (19.7)
Heart failure or reduced LVEF	22 329 (30.2)	50 480 (34.5)	11 650 (28.1)	33 422 (33.0)	3 052 (34.5)	3 022 (34.2)	0.007	3 235 (35.5)	3 216 (35.4)
Diabetes	20 104 (27.2)	40 103 (27.4)	11 630 (28.0)	28 496 (28.2)	2 243 (25.4)	2 275 (25.7)	0.008	2 284 (25.0)	2 263 (24.9)
Hypertension req. treatment	52 406 (71.0)	105 097 (71.8)	31 780 (76.6)	76 923 (76.0)	7 662 (86.6)	7 669 (86.7)	0.002	7 962 (87.3)	7 954 (87.6)
$CHADS_2$ score, mean ± SD	2.4 ± 1.5	2.4 ± 1.4	2.5 ± 1.3	2.5 ± 1.2	2.1 ± 1.1	2.1 ± 1.1	0.003	2.1 ± 1.1	2.1 ± 1.1
$CHADS_2$ =0	6 494 (8.8)	10 240 (7.0)	134 (0.3)	356 (0.4)	52 (0.6)	55 (0.6)	0.004	54 (0.6)	58 (0.6)
$CHADS_2$ =1	14 860 (20.1)	28 124 (19.2)	10 602 (25.6)	23 539 (23.3)	2 971 (33.6)	2 912 (32.9)	0.014	3 046 (33.4)	3 025 (33.3)
$CHADS_2$ =2	19 844 (26.9)	43 294 (29.6)	12 969 (31.3)	32 980 (32.6)	3 157 (35.7)	3 239 (36.6)	0.019	3 262 (35.8)	3 254 (35.8)
$CHADS_2$ ≥3	32 645 (44.2)	64 674 (44.2)	17 783	44 284 (43.8)	2666 (30.1)	2640 (29.8)	0.006	2 758 (30.2)	2 744 (30.2)

			(42.9)						
Medications at index date									
ACE inhibitor or ARB	34 899 (47.3)	82 841 (56.6)	21 656 (52.2)	61 435 (60.7)	5529 (62.5)	5573 (63.0)	0.010	6 464 (70.9)	6 368 (70.1)
Amiodarone	1 903 (2.6)	4 859 (3.3)	961 (2.3)	3 259 (3.2)	336 (3.8)	322 (3.6)	0.008	1 009 (11.1)	1 042 (11.5)
Beta-blocker	46 173 (62.5)	88 274 (60.3)	25 990 (62.6)	62 016 (61.3)	6083 (68.8)	6031 (68.2)	0.013	5 797 (63.6)	5 685 (62.6)
Aspirin	5209 (7.1%)	10833 (7.4%)	2 612 (6.3)	6 429 (6.4)	514 (5.8)	557 (6.3)	0.020	2 859 (31.3)	2 773 (30.5)
Clopidogrel	2697 (3.7%)	3697 (2.5%)	1 238 (3.0)	2 177 (2.2)	229 (2.6)	215 (2.4)	0.010	170 (1.9)	168 (1.9)
Digoxin	9 771 (13.2)	33 342 (22.8)	5 147 (12.4)	23 322 (23.1)	1232 (13.9)	1244 (14.1)	0.004	2 916 (32.0)	2 912 (32.1)
Calcium channel blocker	19 659 (26.6)	39 909 (27.3)	12 522 (30.2)	30 379 (30.0)	2965 (33.5)	2994 (33.8)	0.007	2 744 (30.1)	2 823 (31.1)
Statin	39 027 (52.9)	82 086 (56.1)	23 035 (55.5)	58 647 (58.0)	5230 (59.1)	5228 (59.1)	0.000	4 104 (45.0)	4 095 (45.1)
Non-steroidal anti-inflammatory	4 953 (6.7)	8 107 (5.5)	2 939 (7.1)	5 891 (5.8)	487 (5.5)	479 (5.4)	0.004	752 (8.2)	768 (8.5)
Gastric antacid drugs	1 833 (2.5)	3 290 (2.2)	1 042 (2.5)	2 346 (2.3)	180 (2.0)	180 (2.0)	0.000	1 683 (18.5)	1 667 (18.4)
Proton pump inhibitor	2844 (38.0)	47 838 (32.7)	15 197 (36.6)	31 769 (31.4)	3052 (34.5)	3104 (35.1)	0.012		
H2 receptor antagonist	3 188 (4.3)	4 837 (3.3)	1 586 (3.8)	3 006 (3.0)	281 (3.2)	250 (2.8)	0.021		
Renal function, creatinine clearance									
Normal, >80 ml/min	21 591 (29.2)	45 793 (31.3)	12 261 (29.6)	31 451 (31.1)	4 098 (46.3)	4 074 (46.1)	0.005	3 761 (41.2)	3 757 (41.4)
Mild imp., >50 to 80 ml/min	28 976 (39.2)	56 742 (38.8)	17 494 (42.2)	41 290 (40.8)	3 307 (37.4)	3 292 (37.2)	0.004	3 817 (41.9)	3 770 (41.5)
Moderate imp. (>30 to 50 ml/min)	17 007 (23.0)	32 881 (22.5)	9 708 (23.4)	23 316 (23.0)	1 276 (14.4)	1 306 (14.8)	0.010	1 365 (15.0)	1 382 (15.2)
Severe imp. (≤30 ml/min)	4 317 (5.8)	9 251 (6.3)	1 053 (2.5)	4 251 (4.2)	126 (1.4)	132 (1.5)	0.006	137 (1.5)	133 (1.5)
Not reported	1 952 (2.6)	1 665 (1.1)	972 (2.3)	851 (0.8)	39 (0.4)	42 (0.5)	0.005	40 (0.4)	39 (0.4)
Peripheral artery disease	5 984 (8.1)	12 764 (8.7)	2 770 (6.7)	7 516 (7.4)	552 (6.2)	538 (6.1)	0.007		
Aortic plaque	17 919 (24.3)	40 415 (27.6)	8 974 (21.6)	25 193 (24.9)	2 097 (23.7)	2 057 (23.3)	0.011		
Smoking status									
Non-smoker	27 568 (37.3)	51 612 (35.3)	15 949 (38.4)	36 338 (35.9)	3 186 (36.0)	3 164 (35.8)	0.005		
Ex-smoker	40 815 (55.3)	84 850 (58.0)	22 757 (54.9)	58 669 (58.0)	4 925 (55.7)	4 945 (55.9)	0.005		

Current smoker	5 236 (7.1)	9 658 (6.6)	2 688 (6.5)	6 049 (6.0)	735 (8.3)	737 (8.3)	0.001		
Missing	224	211	94	102	0	0			
Alcohol consumption									
Non-drinker	27 185 (36.8)	52 744 (36.0)	14 957 (36.1)	35 905 (35.5)	2 802 (31.7)	2 842 (32.1)	0.010		
Light, 1 to 14 units per week	32 190 (43.6)	66 072 (45.2)	18 762 (45.2)	46 876 (46.3)	4 135 (46.7)	4 153 (46.9)	0.004		
Moderate, 15 to 42 units per week	8 950 (12.1)	15 916 (10.9)	5 053 (12.2)	11 109 (11.0)	1 563 (17.7)	1 515 (17.1)	0.014		
Heavy, > 42 units per week	1 488 (2.0)	2 028 (1.4)	617 (1.5)	1 149 (1.1)	203 (2.3)	204 (2.3)	0.001		
Missing	3 901	9 223	2 032	5 893	143	132			
Socioeconomic status (England IMD2015 Quintile)									
1 (least deprived)	18 893 (25.6)	36 046 (24.6)	10 867 (26.2)	25 270 (25.0)	2 246 (25.4)	2 231 (25.2)	0.004		
2	17 203 (23.3)	33 585 (23.0)	9 768 (23.5)	23 473 (23.2)	2 098 (23.7)	2 057 (23.3)	0.011		
3	14 591 (19.8)	29 856 (20.4)	8 207 (19.8)	20 704 (20.5)	1 715 (19.4)	1 759 (19.9)	0.013		
4	12 283 (16.6)	25 614 (17.5)	6 767 (16.3)	17 498 (17.3)	1 443 (16.3)	1 465 (16.6)	0.007		
5 (most deprived)	10 804 (14.6)	21 066 (14.4)	5 843 (14.1)	14 098 (13.9)	1 344 (15.2)	1 334 (15.1)	0.003		
Missing	69	165	36	116	0	0			
Ethnicity									
White	70 703 (95.7)	141 019 (96.4)	39 685 (95.7)	97 735 (96.6)	8 424 (95.2)	8 444 (95.5)	0.011	7 536 (82.6)	7 493 (82.5)
Black	714 (1.0)	1 326 (0.9)	372 (0.9)	821 (0.8)	104 (1.2)	103 (1.2)	0.001	125 (1.4)	102 (1.1)
Asian	1 371 (1.9)	2 481 (1.7)	774 (1.9)	1 536 (1.5)	214 (2.4)	209 (2.4)	0.000	1 310 (14.4)	1 332 (14.7)
Other	198 (0.3)	356 (0.2)	113 (0.3)	232 (0.2)	22 (0.2)	22 (0.2)	0.000	149 (1.6)	153 (1.7)
Mixed	152 (0.2)	308 (0.2)	75 (0.2)	190 (0.2)	25 (0.3)	28 (0.3)	0.006	0	0
Unknown	385 (0.5)	448 (0.3)	252 (0.6)	350 (0.3)	42 (0.5)	25 (0.3)	0.031	0	0
Charlson comorbidity components									
Chronic obstructive pulmonary disease	10 324 (14.0)	19 033 (13.0)	5 411 (13.0)	12 573 (12.4)	1 138 (12.9)	1 141 (12.9)	0.001		
Connective tissue disease	5 377 (7.3)	9 784 (6.7)	3 000 (7.2)	6 744 (6.7)	536 (6.1)	534 (6.0)	0.001		
Peptic ulcer	4 400 (6.0)	8 399 (5.7)	2 161 (5.2)	5 458 (5.4)	411 (4.6)	393 (4.4)	0.010		

Liver disease	761 (1.0)	1 291 (0.9)	263 (0.6)	642 (0.6)	76 (0.9)	61 (0.7)	0.019		
Hemiplegia	265 (0.4)	559 (0.4)	147 (0.4)	379 (0.4)	24 (0.3)	16 (0.2)	0.019		
Non-haematological Cancer	12 567 (17.0)	23 383 (16.0)	6 019 (14.5)	14 413 (14.2)	1 066 (12.1)	1 146 (13.0)	0.027		
Haematological cancer	1 966 (2.7)	3 481 (2.4)	951 (2.3)	2 231 (2.2)	174 (2.0)	163 (1.8)	0.009		
BMI - kg/m^2, median (IQR)	28 (24, 32)	28 (23, 32)	28 (25, 32)	28 (25, 32)	29 (26, 33)	29 (26, 33)	0.003		
Missing	2 270	5 858	1 166	3 593	0	0			

Table 3: Baseline characteristics of patients with Atrial Fibrillation prescribed apixaban and warfarin in CPRD Aurum compared with ARISTOTLE participants: i) before and ii) after applying ARISTOTLE inclusion and exclusion criteria and iii) after matching to the trial participants.

ACE = angiotensin-converting enzyme; ARB = angiotensin-receptor blocker; BMI=body mass index; $CHADS_2$ = stroke risk factor score based on Congestive heart failure, Hypertension, Age ≥ 75 years, Diabetes, prior Stroke; CPRD = Clinical Practice Research Datalink; IMD2015= Index of Multiple Deprivation 2015; imp.=impairment; IQR=interquartile range; LVEF=left ventricular ejection fraction; PSM = propensity score matching; SD=standard deviation; SE=systemic embolism; TIA=transient ischemic attack; VKA = vitamin K antagonist;

Main results

The hazard ratio (HR) for stroke/systemic embolism (SE) in the propensity score matched groups was 0.98 (95% CI 0.82,1.19) (Figure 3 and Table A3 in S3 Appendix). This association was consistent with the non-inferiority margin (upper limit of the 95% CI less than 1.38) [7] but did not show superiority as predicted by ARISTOTLE [HR 0.79 (95% CI 0.66,0.95)] (Figure 3 and Appendix Table A2 in S3 Appendix). The outcome of all-cause mortality also showed non-inferiority [Aurum 1.03 (0.93,1.14) vs trial 0.89 (0.80,0.998)] but did not indicate apixaban superiority. Absolute event rates for the primary outcome and components were close to the trial for apixaban – for example [comparing Aurum vs trial] stroke/SE event rate of 1.27%/year vs. 1.27% whereas the warfarin group had a lower event rate compared with ARISTOTLE (stroke/SE event rate of 1.29%/year vs. 1.60% and haemorrhagic stroke 0.33 %/yr vs 0.47%/yr) (Figure 3). Mean duration of follow-up in the cohort was 1.8 years in the apixaban arm and 2.2 years in the warfarin arm.

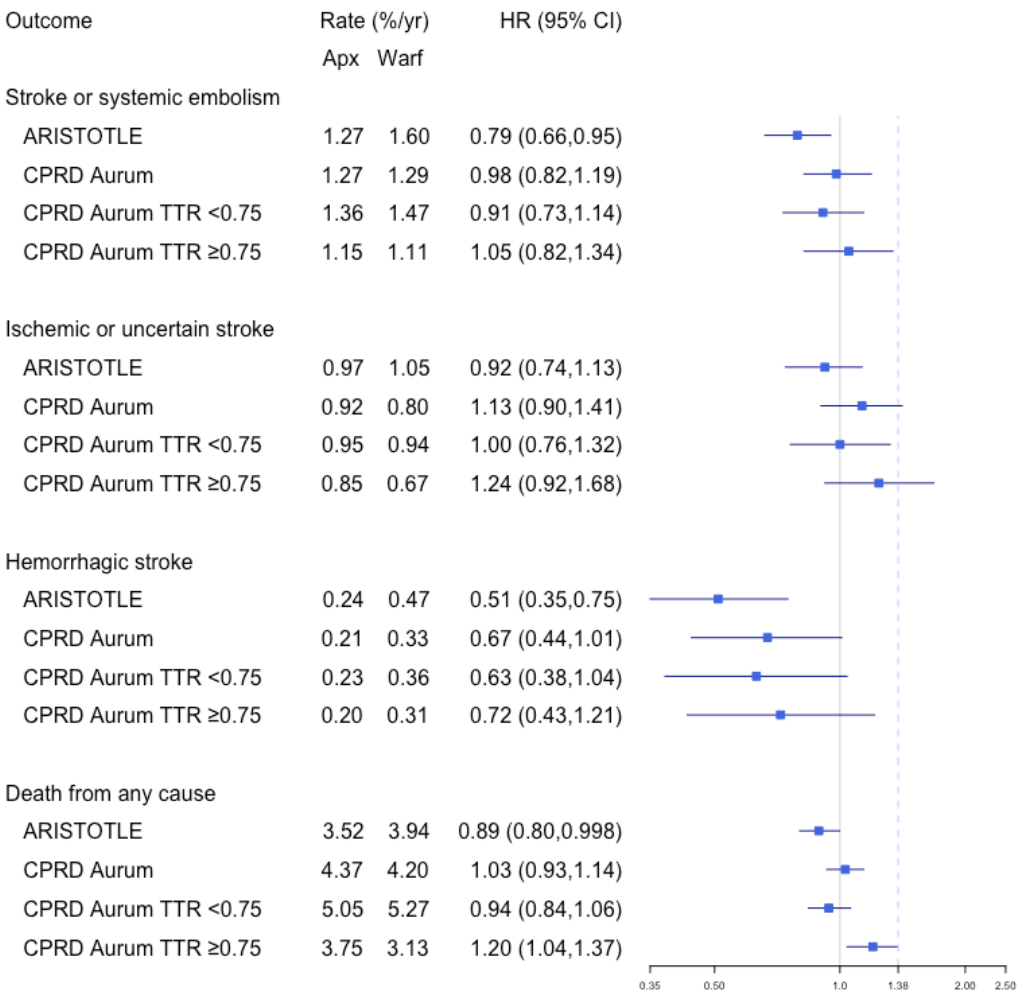


Figure 3: Forest plot showing hazard ratios (dots) and 95% confidence intervals (lines) for apixaban vs warfarin. Absolute event rates (%/year) and Hazard Ratio (95% Confidence Intervals) are presented for key effectiveness outcomes in i) ARISTOTLE, and ii) CPRD Aurum trial-matched cohort, iii) CPRD Aurum trial-matched with TTR<0.75, and iv) CPRD Aurum trial-matched with TTR≥0.75.
Dashed line shows non-inferiority margin 1.38 for the upper bound of the 95% CI of the hazard ratio used in ARISTOTLE for the primary outcome of stroke or systemic embolism.
For the analysis by TTR inverse probability of treatment weighting was applied to the apixaban users targeting the treatment effect in the warfarin users with TTR <0.75 and TTR ≥0.75.
CI=confidence interval; CPRD=Clinical Practice Research Datalink; HR=hazard ratio; TTR=time in therapeutic range.

Analysis of impact of warfarin time in therapeutic range (TTR)

TTR was higher in the CPRD cohort than in ARISTOTLE (mean 0.73 vs. 0.62, median 0.76 vs 0.66). Analysis by TTR suggested non-inferiority of apixaban vs warfarin in those with TTR < 0.75 [stroke/SE 0.91 (0.73,1.14), all-cause death 0.94 (0.84,1.06)] (Figure 3).

Apixaban was associated with similar hazards for stroke by category of TTR and increased

hazards of death compared to warfarin in those with well-controlled warfarin treatment only (TTR ≥ 0.75) [stroke/SE 1.05 (0.82, 1.34), all-cause death 1.20 (1.04,1.37)] (Figure 3).

Analysis of apixaban dose-adjustment

The proportion of patients meeting the criteria for reduced dose apixaban (Table 4) was similar between the CPRD ARISTOTLE-analogous apixaban, warfarin, and RCT apixaban groups (4.9%, 4.9%, and 4.7% respectively). When including the additional NICE criteria of creatinine clearance 5.1% of apixaban users in the ARISTOTLE-analogous cohort had an indication for reduced-dose apixaban yet a larger proportion (14.3%) were prescribed reduced dose apixaban implying some patients in CPRD Aurum may have been prescribed the wrong dose and/or information on criteria for dose reduction may have been missing from CPRD Aurum.

Table 4: Apixaban Dose-adjustment in CPRD Aurum compared with ARISTOTLE

	CPRD Aurum ARISTOTLE-analogous Apixaban (N = 8846)	CPRD Aurum ARISTOTLE-analogous Warfarin (N = 8846)	ARISTOTLE RCT Apixaban (N=9120)
Standard 5.0 mg BID dose	7580 (85.7%)	N/A	8692 (95.3%)
Reduced 2.5mg BID dose	1266 (14.3%)	N/A	428 (4.7%)
Reduced dose indicated per ARISTOTLE criteria	434 (4.9%)	436 (4.9%)	428 (4.7%)
Reduced dose indicated per NICE criteria	454 (5.1%)	459 (5.2%)	NR

NICE criteria for dose-adjustment included additional criteria of creatinine clearance 15–29 mL/minute. N/A=Not applicable. NR=Not reported.

A further analysis of the quality of dose-adjustment in patients in CPRD Aurum (Table 5) indicated 10.5% of patients may have been prescribed an incorrect dose of apixaban at the index prescription based on the data contained in their EHRs. The majority of incorrect dose relating to patients being prescribed reduced-dose apixaban despite not meeting the criteria for dose reduction. A large proportion of patients prescribed an incorrect dose had only 1

dose adjustment criteria (59.6% of those with incorrect dose) suggesting some prescribers may have thought a dose reduction was warranted when only 1 criteria was present. Other possible reasons for the incorrect dose-adjustment observed here may be data on the criteria missing from the EHR record (ie incorrect ascertainment) or consideration of other medical history which made a prescriber adjust the dose.

Table 5: Quality of apixaban dose-adjustment in CPRD Aurum ARISTOTLE-analogous cohort

Dose Status Against NICE Criteria For Dose-adjustment at Index Date	CPRD Aurum ARISTOTLE-analogous Apixaban (N = 8846)
Patients on correct dose	7921 (89.5%)
Patients on incorrect dose	925 (10.5%)
Standard 5.0 mg BID dose despite meeting criteria for dose reduction	59 (0.7%)
Reduced 2.5mg BID dose despite not meeting criteria for dose reduction	866 (9.8%)
0 dose adjustment criteria recorded in EHR	313 (3.5%)
1 dose adjustment criteria recorded in EHR	553 (6.3%)
Age > 80 years	389 (4.4%)
Body weight ≤ 60 kg	57 (0.6%)
Serum creatinine ≥ 1.5 mg/dL	107 (1.2%)

To assess the impact of the quality of dose-adjustment in the CPRD cohort on the effectiveness of apixaban a supplementary post hoc analysis was performed looking at the results in the subset of apixaban patients prescribed the correct dose (N=7921) compared with IPTW re-balanced warfarin comparators. The results in this subset were consistent with the primary results showing apixaban to be non-inferior to warfarin (Stroke/SE 0.96 [0.78,1.17], death 0.97 [0.87,1.09]) with the results moving slightly closer to those observed in ARISTOTLE.

Safety results

The analysis for safety outcomes is presented in Figure 4 and Table A5 in S3 Appendix; patients on apixaban had a lower risk of major bleeding compared with those on warfarin, HR (95% CI) 0.88 (0.77,1.00), consistent with ARISTOTLE. Analysis by TTR suggested

superiority of apixaban for major bleeding in those with TTR <0.75 [0.74 (0.63,0.86)] whereas apixaban users had a similar risk of major bleeding compared with those with optimal warfarin control (TTR ≥ 0.75) [1.08 (0.90,1.30)].

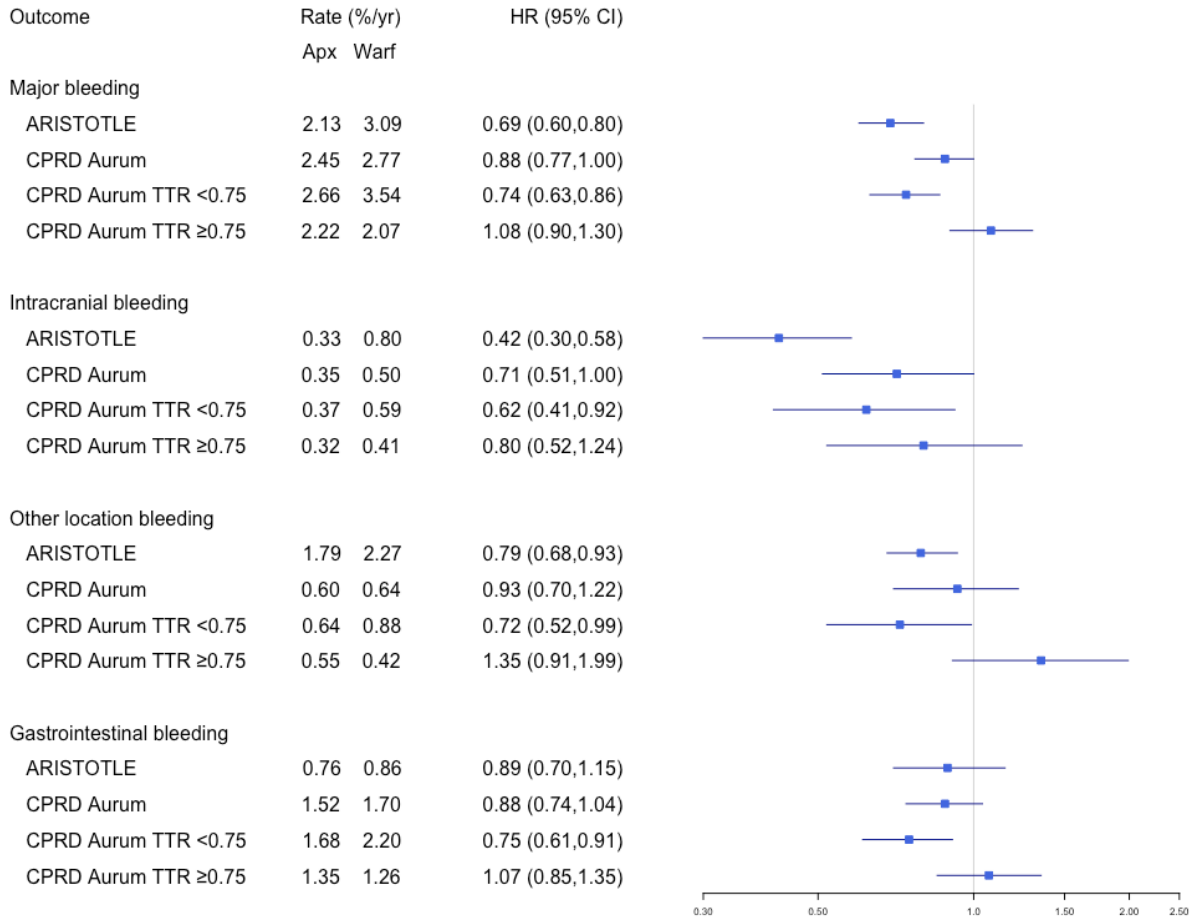


Figure 4: Forest plot showing hazard ratios (dots) and 95% confidence intervals (lines) for apixaban vs warfarin. Absolute event rates (%/year) and Hazard Ratio (95% Confidence Intervals) are presented for key safety outcomes in i) ARISTOTLE, and ii) CPRD Aurum trial-matched cohort, , iii) CPRD Aurum trial-matched with TTR<0.75, and iv) CPRD Aurum trial-matched with TTR≥0.75. For the analysis by TTR inverse probability of treatment weighting was applied to the apixaban users targeting the treatment effect in the warfarin users with TTR <0.75 and TTR ≥0.75.
CI=confidence interval; CPRD=Clinical Practice Research Datalink; HR=hazard ratio; TTR=time in therapeutic range.

Sensitivity analyses

Table A7 in S3 Appendix shows the proportion of patients switching treatment. A higher proportion of patients on warfarin switched to an alternative OAC during follow-up compared with those on apixaban (16.3% vs 6.1%).

Comparing patients who switched treatment during follow-up with those that continued on index treatment (Table A8 in S3Appendix) suggests possible selection bias due to attrition in on-treatment analyses with median TTR markedly lower in warfarin users who switched treatments compared with persistent warfarin users (median TTR 0.64 vs 0.78). On-treatment analyses would likely be biased against apixaban since patients doing badly on warfarin (i.e. with low TTR) who would be more likely to experience events in the warfarin arm would be censored at treatment switch.

On-treatment analyses censoring around treatment switch or discontinuation are presented for the effectiveness analyses in the appendix (Table A6 in S3 Appendix); the results show evidence of the expected attrition bias against apixaban when compared with the ITT results in Figure 2, for example HR for stroke/SE is 1.04 (0.86, 1.25) in the on-treatment compared with 0.98 (95% CI 0.82, 1.19) in the ITT analysis.

Repeating the analysis with start of study period shifted forwards a year to investigate the impact of inclusion of early adopters yielded similar results to the primary analysis (Table A9 in S3 Appendix).

Prior INR control was not included in the propensity score models for the VKA-experienced due to a high rate of missing prior INR data (missing for 34% in the apixaban arm). A post-hoc sensitivity analysis including a prior INR control variable in the PSM gave results consistent with the primary results [Stroke/SE HR 95%CI 1.02 (0.86,1.21)]. Details of this post hoc analysis are in S3 Appendix.

Discussion

In our emulation of ARISTOTLE using UK routinely-collected healthcare data we found results that met our predefined criteria for comparability with the trial. We saw non-inferiority of apixaban vs warfarin for prevention of stroke or systemic embolism, all-cause mortality, and major bleeding, but did not see superiority of apixaban vs warfarin for these outcomes as was seen in ARISTOTLE. We found higher TTR in the patients using warfarin in our cohort compared with the warfarin arm of ARISTOTLE (median 0.76 vs 0.66). While our analysis by TTR showed non-inferiority of apixaban vs warfarin for our stroke or systemic embolism outcome, we observed an increased risk of death on apixaban compared with patients well-controlled on warfarin (TTR $\geq$ 0.75) but not when compared with those on poorly controlled warfarin (TTR<0.75). For major bleeding, while apixaban was superior when compared to those on poorly controlled warfarin, there was no difference when compared to those on well controlled warfarin. We saw evidence suggesting sub-optimal dosing of apixaban in our cohort with approximately 10% of patients in the apixaban arm prescribed the reduced dose without meeting the criteria for the reduced dose.

We found the differences in the overall treatment-effect estimates between our cohort and ARISTOTLE may be explained by: the lower proportion of Asian patients in our cohort, differences in INR control in the warfarin arm of our cohort compared with ARISTOTLE, and the higher proportion of patients prescribed a reduced dose of apixaban in our cohort compared with ARISTOTLE.

Our findings are consistent with a UK study of ischemic stroke which compared DOACs with warfarin [29]. A Danish study found similar results to ours for stroke/SE [30] although they found apixaban users had a lower risk of death, a study of US claims data [31] also found apixaban was associated with a lower risk of death. A systematic review and meta-analysis of

16 studies [32] found pooled results for stroke and ICH that were consistent with ours. One study (in US claims data) also aimed to replicate ARISTOTLE [33, 34] and in contrast to our study found superiority for apixaban for stroke/SE, which may be linked to population differences such as lower TTR in US patients on warfarin [35] and differences in ethnicity. None of these studies matched to the ARISTOTLE trial participants, included prevalent users. Further details on these studies including design and key results are summarized in Table A10 in S3 Appendix.

A key strength of our study was the use of a framework which sampled prevalent users (the continuing users of warfarin in this study) in a way that avoided selection bias facilitating the construction of a cohort of patients similar to the target trial population, which included both new users of apixaban and warfarin (VKA-naïve) and patients with prior VKA exposure (VKA-experienced) that were randomised to stay on warfarin or switch to apixaban. The use of propensity score matching, stratified by treatment history, enabled us to select a matched cohort well balanced on important covariates. The successful emulation of ARISTOTLE by our study shows that valid treatment effects can be obtained for important outcomes with OACs using non-interventional methods with routinely collected clinical data. Having validated this framework, in future studies we can look at the effectiveness of oral anticoagulants in AF patient groups not included or underrepresented in the RCT such as elderly patients and those at increased bleeding risk. We also recommend future analyses with an extended follow-up period compared with this study to compare the long term outcomes seen in the non-interventional cohort with projected long-term outcomes from the RCT.

An additional strength of our study was the ability to explore the quality of warfarin treatment in our cohort and the impact of INR control on the treatment effect estimates. Our finding that the benefits of apixaban vs warfarin for some outcomes depended on the quality

of INR control in the warfarin arm answers questions raised in the NICE premeeting briefing which looked at apixaban in the NVAF population and noted the TTR seen in ARISTOTLE "may be lower than what is typical in UK clinical practice" and "apixaban compared with well-controlled warfarin (TTR 75% or more) may not be superior in the long term" [8]. ARISTOTLE presented outcomes by centre (for example hospital) TTR quartile and did not show a signal of treatment efficacy differing by centre TTR quartile. We were able to use inverse probability of treatment weighting to estimate the treatment effect in the different warfarin TTR groups and used predicted TTR for warfarin users missing TTR to attempt to limit the risk of selection bias.

Whilst our study aimed to emulate ARISTOTLE using suitable methods there were several limitations. Some of the criteria assessed for ARISTOTLE eligibility may not be well recorded in CPRD leading to a risk of misclassification. Furthermore, misclassification of ARISTOTLE eligibility criteria and baseline covariates could be differential by treatment in the VKA-experienced patients if criteria such as renal function are more likely to be checked before changing treatment. However, the most important risk factors for the primary outcome of stroke (the components of CHA_2DS_2-VASc stroke risk score) are mostly well recorded in CPRD Aurum and HES.

Our cohort did not attempt to match the trial on the use of concomitant medications in order for our cohort to reflect typical UK prescribing. In ARISTOTLE 31% of participants were using aspirin and 11% using amiodarone at baseline whereas, in our cohort only 6% were recorded as using aspirin and 4% amiodarone. Amiodarone potentiates the effects of warfarin and concomitant use of amiodarone with DOACs is associated with increased risk of major bleeding [36], whilst concomitant use of aspirin increases the risk of bleeding for both warfarin[37] and DOACs [38]. The difference in concomitant medication usage between our

cohort and the trial population may explain some of the observed differences in treatment effects.

A key limitation of our study was the inability to match ARISTOTLE on ethnicity meaning the CPRD Aurum cohort included a low number of patients from Asian and Hispanic groups when compared with the RCT (14.5% of participants in ARISTOTLE were Asian compared to 2.4% in our ARISTOTLE-analogous CPRD cohort). There are known racial differences in the treatment effects of OACs with Asian patients experiencing a higher risk of haemorrhagic stroke and intracranial haemorrhage compared with White patients; in ARISTOTLE Asian participants experienced double the risk of stroke or systemic embolism when on warfarin therapy when compared with white participants [39]. The reasons for the increased risk of bleeding associated with warfarin therapy in Asian patients is hypothesised to be associated with differences in drug metabolism and prevalence of cerebral microbleeds [40]. The difference in proportion of Asian patients between our cohort and ARISTOTLE is therefore likely to explain some of the differences in treatment effects seen and limits the generalisability of our study, with the results of our study of most relevance to White patients. This limitation on ethnicity arose from the data source used and time period studied (patients with AF in CPRD Aurum 2013-2019) which had a low proportion of Asian patients, likely due to AF being associated with older age combined with a lower prevalence of AF in Asian patients compared with white patients [41]. Whilst CPRD Aurum is largely representative of the UK population in relation to ethnicity [42], diversity is still limited for older individuals. Despite this, CPRD Aurum has shown to be a useful resource for investigating treatment effects in different ethnic groups for indications such as hypertension which is more prevalent and occurs at a younger age in ethnic minority groups, with similar trial replication methods used to compare antihypertensive treatment effects in underrepresented ethnic groups [13].

The approach our study used for handling missing data on baseline covariates relied on assumptions on the relationship between missingness, treatment, and outcomes which may not be valid; however the low proportion of missing data means that this is unlikely to have impacted the results. In the coarsened exact matching step the choice of variables will have an impact on the resulting cohort selected meaning a different combination of variables could lead to different results. There is a risk that residual confounding may be present despite the use of propensity score matching. The use of propensity score matching also has the potential to introduce bias by dropping patients from the cohort [19], however propensity score matching is well suited to the process of trial emulation including prevalent users and a low number of apixaban users were dropped due to unsuccessful matching. The inclusion of prevalent users of warfarin in the cohort risks the introduction of selection bias[20,21]; this was avoided by use of a method shown to produce unbiased estimates in a simulation study [21]. We found consistent results between our new and prevalent user strata across multiple outcomes providing reassurance the method used was likely to have successfully avoided selection bias.

Apixaban along with other DOACs were rapidly adopted as preferred first line OAC in AF during the study period; it was therefore not possible to match on calendar date leading to a difference in follow-up time between the treatment arms in our cohort. A higher proportion of warfarin users switched to alternative OAC during follow-up compared with those prescribed apixaban (16% vs 6%). The impact of this differential switching during follow-up was addressed in the sensitivity analyses. The availability of new alternative treatments during the study period also means there is a risk of channelling bias in that over time the patients still on warfarin are more likely to be those doing well on warfarin. INR control prior to the index date was not included in the propensity score for the prevalent users due to a high rate of missing data, however, other variables associated with poor INR control were included in the

models and an exploratory post-hoc analysis including a variable for poor INR control gave results consistent with the primary results.

Adherence to treatment was difficult to assess in our study due to automatic repeat prescriptions; treatment persistence was more useful in providing a measure of pattern of medicine use over time. In the analysis by TTR the adherence of patients using apixaban was not accounted for, however, a previous UK study showed apixaban had higher adherence than VKAs [41] meaning we would expect to see better effectiveness outcomes in apixaban. Furthermore, the use of IPTW in the analysis by TTR means predictors of poor adherence are likely to have been balanced between treatments. The analysis of TTR is limited by this being a post-baseline measure available for only one treatment arm leading to a risk of selection bias in this analysis – patients with TTR available in the study may be more healthy than those without this measure given that patients have to survive and not be hospitalised to have INR measurements available in CPRD Aurum. The limitation of use of a post-baseline measurement available for one treatment arm was also evident in the RCTs of DOACs vs warfarin and is mitigated in our study through the use of IPTW and predicted TTR for patients that were missing TTR (using a model to predict TTR that used INR measurements restricted to the first year of follow-up). Given the risk of selection bias in the analysis by TTR and risk of miss-classification of TTR for those missing TTR, these results should be considered exploratory and interpreted with caution. Sensitivity analyses in our cohort using an on-treatment censoring scheme showed evidence of attrition bias. The regular measurement of INR and availability of alternative anticoagulants makes warfarin therapy particularly prone to attrition bias since a patient may be more likely to switch to a DOAC if their INR is frequently out of the optimal range or if they have not been adhering to scheduled INR testing.

To conclude, we found that applying a reference trial emulation approach allowed us to emulate a landmark randomized trial of apixaban versus warfarin using UK non-interventional data, with results meeting pre-specified benchmarking criteria based on the reference trial results. This trial emulation method provides valid treatment effect estimates for apixaban compared to warfarin and can be used to determine risks and benefits of AF medications in people treated in routine clinical care. This study demonstrates a successful real world application of novel methods that have been proposed for the inclusion of prevalent users in observational studies, with the application of an adaptation to mimic the screening process making the method suitable for emulation of RCTs that include prevalent users. These methods could be adapted for emulation of RCTs in other therapeutic areas and for looking at patient groups under-represented or excluded from RCTs.

The weaker overall treatment benefit observed in our cohort appears to be due to a higher proportion of patients with well-controlled warfarin in the UK clinical context, compared with the trial. Our exploratory analysis by TTR showed similar results for stroke and a greater benefit for apixaban for major bleeding compared with TTR<0.75; conversely a slightly higher risk of death was observed on apixaban compared with well-controlled warfarin.

The views expressed in this paper are those of the author and not do not necessarily reflect those of the SFDA or its stakeholders. Guaranteeing the accuracy and the validity of the data is a sole responsibility of the research team.

References

[1] 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. doi: 10.1136/heartjnl-2018-312977.

[2] Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. 1991 Aug;22(8):983-8. doi: 10.1161/01.str.22.8.983.

[3] Friberg L, Rosenqvist M, Lindgren A, Terént A, Norrving B, Asplund K. High prevalence of atrial fibrillation among patients with ischemic stroke. **Stroke.** 2014 Sep;45(9):2599–605. doi: 10.1161/STROKEAHA.114.006070.
[4] Ali AN, Abdelhafiz A. Clinical and Economic Implications of AF Related Stroke. J Atr Fibrillation. 2016 Feb. 29;8(5):1279. doi: 10.4022/jafib.1279.
[5] Andrew NE, Thrift AG, Cadilhac DA. The prevalence, impact and economic implications of atrial fibrillation in stroke: what progress has been made? **Neuroepidemiology.** 2013; 40(4):227–39. doi: 10.1159/000343667.
[6] Apostolakis S, Sullivan RM, Olshansky B, Lip GYH. 053 Factors affecting quality of warfarin anticoagulation in patients with atrial fibrillation: insights from affirm. Heart. 2012;98:A32. doi:10.1136/heartjnl-2012-301877b.53.
[7] 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 Sep 15;365(11):981-92, doi: 10.1056/NEJMoa1107039.
[8] European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) Assessment report Eliquis apixaban. Procedure No.:EMEA/H/C/002148/X/04/G. EMA/641505/2012, 2012, page 35. Available from: https://www.ema.europa.eu/en/documents/variation-report/eliquis-h-c-2148-x-0004-g-epar-assessment-report-extension_en.pdf
[9] National Institute for Health and Clinical Excellence. Premeeting briefing. Apixaban for preventing stroke and systemic embolism in people with nonvalvular atrial fibrillation. Issue date: November 2012. Available from: https://www.nice.org.uk/guidance/ta275/documents/stroke-and-systemic-embolism-prevention-nonvalvular-atrial-fibrillation-apixaban-premeeting-briefing2
[10] Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT DUPLICATE Initiative. Circulation. 2021 Mar 9.143(10):1002-1013. doi: 10.1161/CIRCULATIONAHA.120.051718. Epub 2020 Dec 17.
[11] Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, et al. Real world effects of COPD medications: a cohort study with validation against results from randomised controlled trials. Eur Respir J. 2021 Mar 25;57(3):2001586. doi: 10.1183/13993003.01586-2020.
[12] Dahabreh IJ, Robins JM, Hernán MA. Benchmarking Observational Methods by Comparing Randomized Trials and Their Emulations. Epidemiology. 2020 Sep;31(5):614-619. doi: 10.1097/EDE.0000000000001231.
[13] Baptiste PJ, Wong A YS, Schultze A, Clase CM, Leyrat C, Williamson E, et al. Comparative effectiveness of ARB and ACEi for cardiovascular outcomes and risk of angioedema among different ethnic groups in England: an analysis in the UK Clinical Practice Research Datalink with emulation of

a reference trial (ONTARGET). medRxiv preprint 2024 Jan 17 doi: doi:10.1101/2024.01.17.24301397.
[14] Powell EM, Douglas IJ, Gungabissoon U, Smeeth L, Wing K. Real-world effects of medications for stroke prevention in atrial fibrillation: protocol for a UK population-based non-interventional cohort study with validation against randomized trial results. BMJ Open 2021;11:e042947. doi:10.1136/bmjopen-2020-04294.
[15] Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. Int J Epidemiol. 2019 Dec 1;48(6):1740-1740g. doi: 10.1093/ije/dyz034.
[16] Rose M, Beasley N. Center for Drug Evaluation and Research Application number: 202155Orig1s000 MEDICAL REVIEW(S). Available for download at: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/202155Orig1s000TOC.cfm
[17] Vinereanu D, Stevens SR, Alexander JH, Al-Khatib SM, Avezum A, Bahit MC, et al. Clinical outcomes in patients with atrial fibrillation according to sex during anticoagulation with apixaban or warfarin: a secondary analysis of a randomized controlled trial. Eur Heart J. 2015 Dec 7;36(46):3268-75. doi: 10.1093/eurheartj/ehv447.
[18] Halvorsen S, Atar D, Yang H, De Caterina R, Erol C, et al. Efficacy and safety of apixaban compared with warfarin according to age for stroke prevention in atrial fibrillation: observations from the ARISTOTLE trial. Eur Heart J. 2014 Jul 21;35(28):1864-72. doi: 10.1093/eurheartj/ehu046.
[19] Iacus SM, King G, Porro G. Causal Inference without Balance Checking: Coarsened Exact Matching Political Analysis. 2012;20(1):1-24. doi:10.1093/pan/mpr013
[20] Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. Pharmacoepidemiol Drug Saf. 2017 Apr;26(4):459-468. doi: 10.1002/pds.4107.
[21] Webster-Clark M, Mavros P, Garry EM, Stürmer T, Shmuel S, Young J, et al. Alternative analytic and matching approaches for the prevalent new-user design: A simulation study. Pharmacoepidemiol Drug Saf. 2022 Jul;31(7):796-803. doi: 10.1002/pds.5446.
[22] Austin, P.C. (2009), Some methods of propensity-score matching had superior performance to others: results of an empirical Investigation and Monte Carlo simulations. Biom. J. 2009 Feb;51(1): 171-84. doi: 10.1002/bimj.200810488
[23] Woodfield R, Grant I, UK Biobank Stroke Outcomes Group, UK Biobank Follow-Up and Outcomes Working Group, Sudlow CL. Accuracy of Electronic Health Record Data for Identifying Stroke Cases in Large-Scale Epidemiological Studies: A Systematic Review from the UK Biobank Stroke Outcomes Group. PLoS One. 2015 Oct 23;10(10):e0140533. doi: 10.1371/journal.pone.0140533.
[24] Austin P.C, Small DS. The use of bootstrapping when using propensity-score matching without replacement: A simulation study. Stat Med. 2014 Oct 30; 33(24), 4306–19. doi: 10.1002/sim.6276.

[25] Abadie A., Spiess J. Robust Post-Matching Inference. Journal of the American Statistical Association, 117:538, 983-995, doi: 10.1080/01621459.2020.1840383
[26] Lobo FS, Wagner S, Gross CR, Schommer JC. Addressing the issue of channeling bias in observational studies with propensity scores analysis. Res SocialAdm Pharm. 2006 Mar;2(1):143-51. doi: 10.1016/j.sapharm.2005.12.001.
[27] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983 Apr;70(1):41-55. doi: 10.1093/biomet/70.1.41.
[28] Blake HA, Leyrat C, Mansfield KE, Seaman S, Tomlinson LA, Carpenter J, et al. Propensity scores using missingness pattern information: a practical guide. Stat Med. 2020 May 20;39(11):1641–1657. doi: 10.1002/sim.8503.
[29] Vinogradova Y, Coupland C, Hill T, Hippisley-Cox J. Risks and benefits of direct oral anticoagulants versus warfarin in a real world setting: cohort study in primary care. BMJ. 2018 Jul 4; 362:k2505 doi:10.1136/bmj.k2505. Erratum in: BMJ. 2018 Oct 18;363:k4413.
[30] Larsen TB, Skjøth F, Nielsen PB, Kjældgaard 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 Jun 16;353:i3189. doi:10.1136/bmj.i3189.
[31] Li XS, Deitelzweig S, Keshishian A, Hamilton M, Horblyuk R, Gupta K, et al. li Thromb Haemost. 2017 Jun 2;117(6):1072-1082. doi:10.1160/TH17-01-0068
[32] Proietti M, Romanazzi I, Romiti GF, Farcomeni A, Lip GYH. Real-World Use of Apixaban for Stroke Prevention in Atrial Fibrillation: A Systematic Review and Meta-Analysis. Stroke. 2018 Jan;49(1):98-106. doi: 10.1161/STROKEAHA.117.018395.
[33] Franklin J, Brigham and Women's Hospital. Replication of the ARISTOTLE Anticoagulant Trial in Healthcare Claims Data. Available from: https://www.clinicaltrials.gov/ct2/show/NCT04593030
[34] Wang SV, Schneeweiss S, RCT-DUPLICATE initiative, Franklin JM, Desai RJ, Feldman W, et al. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. JAMA. 2023 Apr 25;329(16):1376-1385. doi:10.1001/jama.2023.4221.
[35] Pokorney SD, Simon DN, Thomas L, Fonarow GC, Kowey PR, Chang P, et al. Patients' time in therapeutic range on warfarin among US patients with atrial fibrillation: Results from ORBIT-AF registry. Am Heart J. 2015 Jul;170(1):141-148.e1. doi: 10.1016/j.ahj.2015.03.017.
[36] Chang SH, Chou IJ, Yeh YH, Chiou MJ, Wen MS, Kuo CT, et al. Association Between Use of Non-Vitamin K Oral Anticoagulants With and Without Concurrent Medications and Risk of Major Bleeding in Nonvalvular Atrial Fibrillation. JAMA. 2017 Oct 3;318(13):1250-1259. doi: 10.1001/jama.2017.13883.
[37] Schaefer JK, Li Y, Gu X, et al. Association of Adding Aspirin to Warfarin Therapy Without an Apparent Indication With Bleeding and Other Adverse Events. JAMA Intern Med. 2019;179(4):533–541. doi:10.1001/jamainternmed.2018.7816

[38] Almas T, Musheer A, Ejaz A, et al. Efficacy and safety of direct oral anticoagulants with and without Aspirin: A systematic review and Meta-analysis. Int J Cardiol Heart Vasc. 2022 Mar 26;40:101016. doi: 10.1016/j.ijcha.2022.101016.
[39] Sabir I, Khavandi K, Brownrigg J, Camm AJ. Oral anticoagulants for Asian patients with atrial fibrillation. Nat Rev Cardiol. 2014 May;11(5):290–303 (2014). doi: 10.1038/nrcardio.2014.22.
[40] Kim HK, Tantry US, Smith SC, et al. The East Asian Paradox: An Updated Position Statement on the Challenges to the Current Antithrombotic Strategy in Patients with Cardiovascular Disease. Thromb Haemost. 2021 Apr;121(4):422-432. doi: 10.1055/s-0040-1718729.
[41] Gillott RG, Willan K, Kain K, et al. South Asian ethnicity is associated with a lower prevalence of atrial fibrillation despite greater prevalence of established risk factors: a population-based study in Bradford Metropolitan District. Europace. 2017 Mar 1;19(3):356-363. doi: 10.1093/europace/euw010.
[42] Shiekh, S.I., Harley, M., Ghosh, R.E. et al. Completeness, agreement, and representativeness of ethnicity recording in the United Kingdom's Clinical Practice Research Datalink (CPRD) and linked Hospital Episode Statistics (HES). Popul Health Metrics 21, 3 (2023). https://doi.org/10.1186/s12963-023-00302-0
[43] Banerjee A, Benedetto V, Gichuru P, Burnell J, Antoniou S, Schilling RJ et al. Adherence and persistence to direct oral anticoagulants in atrial fibrillation: a population-based study. Heart. 2020 Jan;106(2):119-126. doi: 10.1136/heartjnl-2019-315307.

Supporting Information

S1: STROBE checklist

S2: ISAC protocol for the ARISTOTLE emulation study

S3: Appendix containing supporting information

Table A1: ARISTOTLE Inclusion and Exclusion Criteria Applied to CPRD Aurum

Table A2: Efficacy Outcomes Results from ARISTOTLE

Table A3: Effectiveness Outcomes Results in the CPRD Aurum ARISTOTLE-analogous Cohort

Table A4: Bleeding Outcomes and Net Clinical Outcomes Results from ARISTOTLE RCT

5.2. Additional results from the emulation of ARISTOTLE in CPRD Aurum

5.2.1. Additional result 1: sensitivity analysis requiring ≥ 2 prescriptions of the index treatment (minimum exposure requirement)

To understand the impact of incorrect ascertainment of exposure, a sensitivity analysis was performed in which patients were required to have at least 2 prescriptions of the treatment of interest. When a patient has only one prescription for the treatment of interest there is uncertainty surrounding whether the patient actually took the treatment.

The results with this minimum exposure requirement (Tables A1.3.1 and A1.3.2 in the appendix) were broadly consistent with the main results for the key outcomes of interest showing non-inferiority of apixaban vs warfarin for Stroke/SE (HR 0.97 95% CI [0.83, 1.13] vs 0.98 [0.82, 1.19] in the main analysis), a similar risk of death (0.99 [0.91, 1.09] vs 1.03 [0.93, 1.14] in the main analysis), and major bleeding (0.85 [0.76, 0.96] vs 0.88 [0.77, 1.00] in the main analysis).

5.2.2. Additional result 2: results in the VKA-naïve and VKA-experienced from the ARISTOTLE emulation

In studies including prevalent and new users it is recommended to compare the results in the new users and prevalent users strata to check results by prior exposure status are plausible and confirm there is no evidence of selection bias. In the case where the selection of continuing users has been conducted in such a way that selection bias is present one may see large implausible treatment effect estimates in the prevalent user group that differ markedly from the new user group. For example, applying a ‘severe comorbid’ exclusion criteria to a pool of potential index dates first before selecting a random index date would enrich the cohort of prevalent users continuing on the comparator with negative outcome events that define the eligibility criteria (such as cancer with short median survival). The effect of this

form of bias may be especially evident in the outcome of all-cause mortality and is a particular risk in trial emulation where the checking of eligibility criteria can introduce bias if an appropriate procedure is not used (such as the prevalent new user design and alternative approaches based on this which should avoid selection bias).

The simulation study by Webster-Clark et al (136) suggested the method used in the ARISTOTLE emulation should not introduce selection bias; the simulation study found the 'forward sampling method' which I used gave a risk ratio of 0.997 for a true underlying effect of 1.00 in the simulated data, a result close to that achieved with the full PNU design (0.999). In the context of VKA exposure previous studies have suggested new users of warfarin have an increased risk of adverse outcomes therefore it was expected that the direction of treatment effect estimates would trend towards slightly more benefit of apixaban over warfarin in the new user strata when compared with the VKA-experienced.

Table 5.1 Results in the CPRD Aurum ARISTOTLE-analogous cohort by prior VKA exposure strata

	Apixaban Event Rate %/yr	Warfarin Event Rate %/yr	Hazard Ratio (95% CI)	P value for interaction
Stroke or systemic embolism				
Prior use of VKA				0.616
No (VKA-naïve)	1.03	1.09	0.92 (0.68, 1.26)	
Yes (VKA-experienced)	1.47	1.44	1.02 (0.81, 1.29)	
Major bleeding				
Prior use of VKA				0.417
No (VKA-naïve)	2.01	2.42	0.81 (0.65, 1.02)	
Yes (VKA-experienced)	2.79	3.04	0.91 (0.77, 1.08)	
All-cause death				
Prior use of VKA				0.914
No (VKA-naïve)	2.94	2.81	1.04 (0.86, 1.26)	
Yes (VKA-experienced)	5.49	5.34	1.03 (0.92, 1.16)	

CI=confidence interval; VKA=vitamin K antagonist; yr= year.

The results by prior VKA experience showed no significant interaction effect (Table S1.3.3) and the results were broadly consistent between the two groups, though there was a trend

towards lower hazard ratios in the VKA-naïve compared to the VKA-experienced in line with the expected trend. ARISTOTLE also saw no significant difference in treatment effect estimates between the VKA-naïve and VKA-experienced.

5.2.3. Additional result 3: supplementary analysis on impact of time in therapeutic range

Warfarin users were dichotomised based on having TTR < 0.75 or TTR ≥0.75. These groups were selected in an attempt to answer a question from a professional group included in the NICE review of ARISTOTLE: "A lower TTR would be associated with more adverse outcomes in the warfarin arm and apixaban compared with well-controlled warfarin (TTR 75% or more) may not be superior in the long term" and also reflects 0.75 being a commonly selected cut-off for defining good INR control in warfarin.

An analysis by TTR was performed targeting i) the treatment effect in apixaban compared with warfarin users with TTR < 0.75 and ii) the treatment effect in apixaban compared with warfarin users with TTR ≥ 0.75 in an attempt to assess the impact of the quality of warfarin therapy on the relative harms and benefits of apixaban vs warfarin.

Analysis conditional on the observation of post-baseline measurements in one treatment arm is likely to result in selection bias - patients on warfarin must survive to have INR measurements in CPRD Aurum enabling derivation of TTR whereas no such restriction would be placed on the patients on apixaban. The minimum time taken for TTR to be derived is immortal time(138) which may bias the hazard ratios upwards making apixaban look less effective than the true underlying treatment effect in an analysis by TTR. In addition to the problem of immortal time bias in the warfarin groups defined by post-index date TTR, additional uncertainty in the analysis by TTR may be caused by the exclusion of patients for whom it is not possible to derive TTR due to missing INR data. An analysis restricted to

those with TTR would be making the assumption that TTR is missing completely at random which is unlikely to be valid. Assumptions made on the patients with missing INR data are likely to have a large impact on the treatment effect estimates and the effect of the selection bias introduced wherein conditioning on presence of a post-index date measure would lead to warfarin users with early events being removed from the analysis and a resulting underestimate of the relative effectiveness of apixaban vs warfarin

The analysis by TTR must therefore attempt to avoid selection bias by including patients with missing TTR. TTR is unlikely to be missing at random for all patients (for example those that experience an event early on may have TTR missing due to hospitalisation and event may be related to lower 'unobserved' TTR), conversely should TTR be missing because a patient is having INR monitored at a specialised clinic or via self-testing then an assumption of missing at random may be more reasonable though a higher TTR than average may also be a reasonable assumption for these reasons for missing data. A pragmatic approach was taken to account for patients with missing TTR by using the data from the patients with TTR data to model TTR based on baseline variables (age, sex, BMI, smoking status, diabetes, congestive heart failure, statins, ACEi or ARB, beta-blockers, digoxin, amiodarone, NSAIDs, PPI, prior VKA exposure (naïve, <6 months prior exposure, >= 6 months prior exposure), alcohol consumption, index of multiple deprivation (IMD)2015_5, renal function, COPD). Use of a model to predict TTR was inspired by the analysis of TTR data in ARISTOTLE(139). A mixed model to predict continuous TTR and a logistic regression model to predict TTR category were trialled and the model with the best performance at predicting TTR in the patients with actual TTR was selected. The logistic regression model had superior performance for classifying patients by TTR category and was used to predict TTR category for the patients on warfarin that were missing TTR, thereby allowing all patients to be included in the analysis and attempt to minimise the risk of selection bias.

Inverse probability of treatment weighting was used estimating propensity scores for the warfarin users in each TTR group of interest compared with all apixaban users in the cohort, and then applying stabilised weights based on these propensity scores to the apixaban users to calculate the average treatment effect in the warfarin users by TTR. When implementing this analysis, a similar approach to the main analysis was used for the calculation of propensity scores: models were fit separately for the new and prevalent users, and in the prevalent users models were fit in the separate treatment history strata.

5.2.4. Additional Result 3: ARISTOTLE-eligible new users

In the literature review in Chapter 3 most non-interventional studies comparing apixaban to warfarin were in new users of OACs and the only other study attempting to emulate ARISTOTLE, by the RCT-DUPLICATE initiative(20), applied only the eligibility criteria, did not match to the trial, and included only new users. To explore the effect of the step of matching to the trial participants on baseline characteristics an additional analysis was conducted using the CPRD Aurum data source omitting this step. Furthermore, to aid comparison with the other non-interventional studies in this therapeutic area (in particular RCT-DUPLICATE), this analysis was restricted to new users.

In this analysis the reference trial eligibility criteria were applied followed by propensity score matching the warfarin and apixaban users. The differences between the two approaches is illustrated in Figure 5.1: the emulation of ARISTOTLE presented in Section 5.1 was looking at the treatment effects in patients in UK data that met the ARISTOTLE eligibility criteria and matched the ARISTOTLE participants on key baseline characteristics including prior VKA exposure. By contrast, this analysis in the ARISTOTLE-eligible new users was looking at treatment effects in patients in UK data that met the ARISTOTLE eligibility criteria and were new users of warfarin or apixaban.

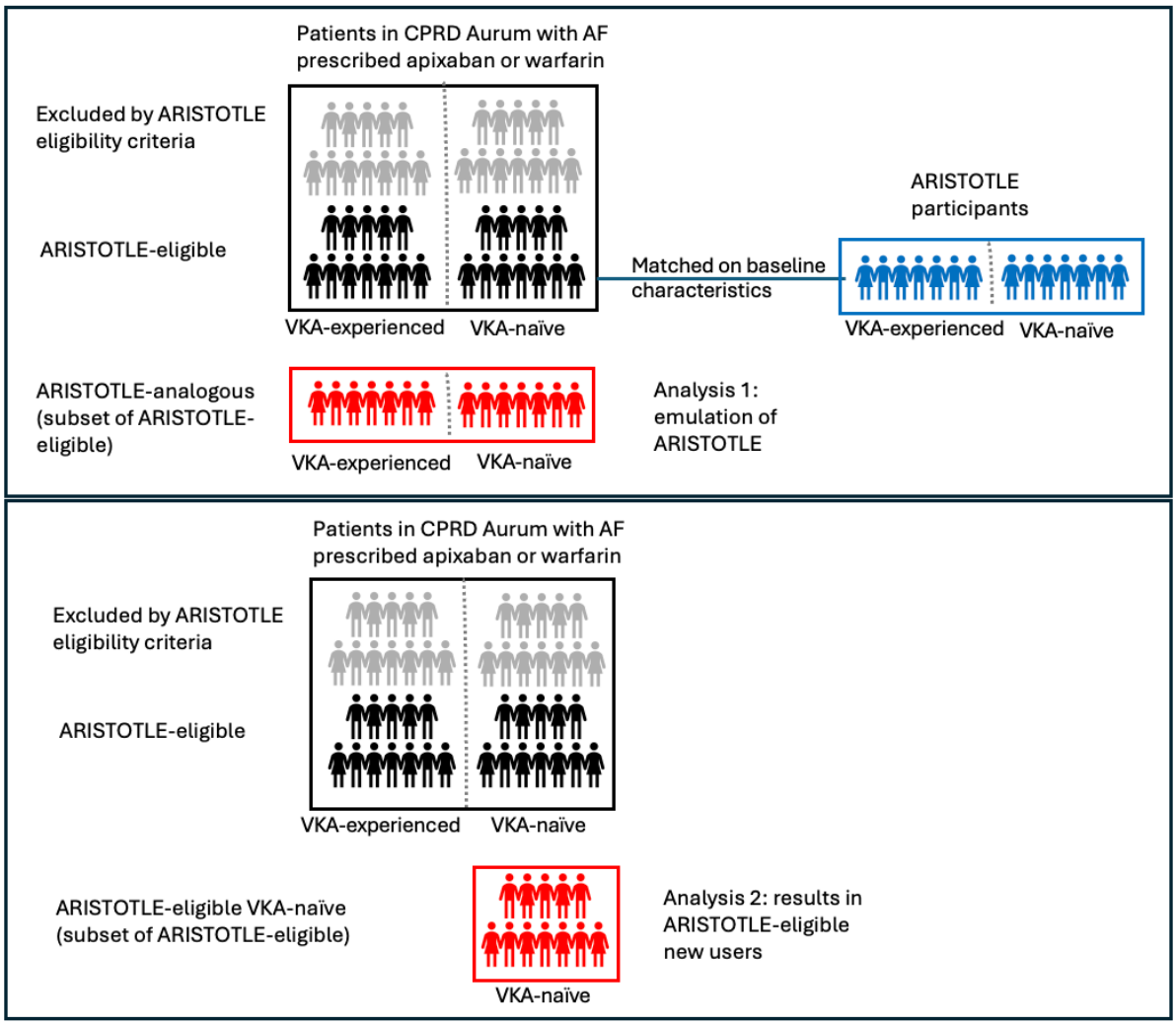


Figure 5.1 Illustration of the different subsets of patients studied in i) Analysis 1 the full emulation of ARISTOTLE and ii) Analysis 2 looking at the ARISTOTLE-eligible new users

AF=atrial fibrillation; VKA=vitamin K antagonist.

Based on the literature review, the comparison of baseline characteristics of patients in CPRD Aurum after applying the ARISTOTLE eligibility criteria, and the results by prior VKA exposure strata, this analysis was expected to include a higher proportion of female and older patients and to show greater benefit of apixaban over warfarin (lower HR estimates) when compared to the ARISTOTLE emulation that included the matching step.

The baseline table of results for the ARISTOTLE-eligible new users in CPRD Aurum (Table 5.2) shows that people in the apixaban arm were older than in the primary emulation that

included the matching step (60.4% aged 75 years or older compared with 31.3%) with a higher proportion of female participants (45.6% vs 35.5%). $CHADS_2$ stroke risk factors also differed in the apixaban arm from the ARISTOTLE emulation with a lower proportion of patients with heart failure or reduced left ventricular ejection fraction (LVEF) (21.5% vs 34.5%), a lower proportion of patients with hypertension requiring antihypertensive treatment (78.2% vs 86.6%), and a higher proportion of patients with $CHADS_2$ score ≥3 (34.1% vs 30.1%). The use of other medications at index date differed from the trial emulation with a higher rate of aspirin use (9.5% vs 5.8% in the apixaban arm) and a lower rate of use of amiodarone and digoxin. A higher proportion of patients had moderate or severe renal impairment compared with the emulation.

The effectiveness results in the ARISTOTLE-eligible matched new user cohort (Table 5.3) showed treatment effect estimates consistent with the ARISTOTLE emulation and the primary effectiveness outcome met the benchmarking criteria. For the safety outcomes (Table 5.4), a lower risk of major bleeding was seen for apixaban vs warfarin (HR 0.86 95% CI [0.78, 0.94]) whereas in the main analysis this trend was borderline significant (HR 0.88 95% CI [0.77, 1.00]), and similarly for intracranial bleeding and other location bleeding. A slightly lower TTR was observed in this analysis compared with the full emulation (mean 0.69 vs 0.73, median 0.74 vs 0.76) which may explain the greater benefit for apixaban over warfarin observed.

Table 5.2 Baseline characteristics of the CPRD Aurum ARISTOTLE-eligible cohort of new users after propensity score matching

Characteristic	Apixaban (N=18 684)	Warfarin (N=18 684)	Standardised difference
Age - years, median (IQR)	77 (70, 82)	77 (70, 83)	.005
Female sex-no.(%)	8529 (45.6)	8486 (45.4)	.005
Systolic blood pressure - mm Hg, median (IQR)	132 (120, 140)	132 (121, 140)	.000
Weight - kg, median (IQR)	80 (69, 94)	81 (70, 94)	.008
Prior myocardial infarction - no. (%)	2159 (11.6)	2160 (11.6)	.000
Prior clinically relevant or spontaneous bleeding – no.(%)	2826 (15.1)	2791 (14.9)	.005
History of fall within previous year – no. (%)	237 (1.3)	218 (1.2)	.009
Qualifying risk factors			
Age ≥ 75 years - no. (%)	11 278 (60.4)	11 270 (60.3)	.001
Prior stroke, TIA, or systemic embolism - no. (%)	3 824 (20.5)	3 858 (20.6)	.005
Heart failure or reduced left ventricular ejection fraction - no. (%)	4 025 (21.5)	3 971 (21.3)	.007
Diabetes - no. (%)	4 954 (26.5)	4 959 (26.5)	.001
Hypertension requiring treatment - no. (%)	14 606 (78.2)	14 581 (78.0)	.003
$CHADS_2$ score			
Mean	2.2 ± 1.2	2.2 ± 1.2	.003
Distribution - no. (%)			
0	50 (0.3)	52 (0.3)	.002
1	5 711 (30.6)	5 710 (30.6)	.000
2	6 557 (35.1)	6 563 (35.1)	.001
≥3	6 366 (34.1)	6 359 (34.0)	.001
Medications at index date - no. (%)			
ACE inhibitor or ARB	10 161 (54.4)	10 258 (54.9)	.010
Amiodarone	343 (1.8)	354 (1.9)	.004
Beta-blocker	11 491 (61.5)	11 438 (61.2)	.006
Aspirin	1 771 (9.5)	1 948 (10.4)	.032
Clopidogrel	710 (3.8)	734 (3.9)	.007
Digoxin	1 622 (8.7)	1 673 (9.0)	.010
Calcium blocker	6 220 (33.3)	6 203 (33.2)	.002
Statin	10 352 (55.4)	10 365 (55.5)	.001
Nonsteroidal anti-inflammatory agent	1 258 (6.7)	1 224 (6.6)	.007
Gastric antacid drugs	471 (2.5)	463 (2.5)	.003
Proton pump inhibitor	6 379 (34.1)	6 371 (34.1)	.001
H2 receptor antagonist	607 (3.2)	611 (3.3)	.001
Renal function, creatine clearance - no. (%)			
Normal, >80 ml/min	6 298 (33.7)	6 342 (33.9)	.005
Mild impairment, >50 to 80 ml/min	8 150 (43.6)	8 124 (43.5)	.003
Moderate impairment (>30 to 50 ml/min)	3 707 (19.8)	3 694 (19.8)	.002
Severe impairment (le 30 ml/min)	392 (2.1)	380 (2.0)	.005
Not reported	137 (0.7)	144 (0.8)	.004
Other risk factors and covariates			
Peripheral artery disease - no. (%)	1 092 (5.8)	1 113 (6.0)	.005
Aortic plaque - no. (%)	3 560 (19.1)	3 598 (19.3)	.005
Smoking status - no. (%)			
Non-smoker	6 884 (36.8)	6 818 (36.5)	.007
Ex-smoker	10 512 (56.3)	10 585 (56.7)	.008
Current smoker	1 288 (6.9)	1 281 (6.9)	.001
Alcohol consumption - no. (%)			
Non-drinker	6 655 (35.6)	6 617 (35.4)	.004
Light drinker, up to 14 units per week	9 325 (49.9)	9 382 (50.2)	.006

Characteristic	Apixaban (N=18 684)	Warfarin (N=18 684)	Standardised difference
Moderate drinker, 15 to 42 units per week	2 424 (13.0)	2 411 (12.9)	.002
Heavy drinker, more than 42 units per week	280 (1.5)	274 (1.5)	.003
Socioeconomic status - no. (%)			
England IMD2015 quintile 1(least deprived)	4 749 (25.4)	4 703 (25.2)	.006
England IMD2015 quintile 2	4 398 (23.5)	4 372 (23.4)	.003
England IMD2015 quintile 3	3 751 (20.1)	3 787 (20.3)	.005
England IMD2015 quintile 4	3 121 (16.7)	3 164 (16.9)	.006
England IMD2015 quintile 5(most deprived)	2 665 (14.3)	2 658 (14.2)	.001
Ethnicity - no. (%)			
White	17 949 (96.1)	17 944 (96.0)	.001
Black	182 (1.0)	180 (1.0)	.001
South Asian	362 (1.9)	360 (1.9)	.001
East Asian	29 (0.2)	31 (0.2)	.003
Mixed	40 (0.2)	43 (0.2)	.003
Other	42 (0.2)	44 (0.2)	.002
Unknown	79 (0.4)	82 (0.4)	.002
Charlson comorbidity index components - no. (%)			
Chronic obstructive pulmonary disease	2228 (11.9)	2196 (11.8)	.005
Connective tissue disease	1152 (6.2)	1152 (6.2)	.000
Peptic ulcer	977 (5.2)	979 (5.2)	.000
Liver disease	95 (0.5)	103 (0.6)	.006
Hemiplegia	44 (0.2)	36 (0.2)	.009
Cancer	2505 (13.4)	2496 (13.4)	.001
Haematological cancer	364 (1.9)	364 (1.9)	.000
BMI - kg/m^2, median (IQR)	28 (25, 32)	28 (25, 32)	.004
ORBIT score			
Low bleeding risk (score 0-2)	13 844 (74.1)	13920 (74.5)	.009
Medium bleeding risk (score 3)	2 659 (14.2)	2570 (13.8)	.014
High bleeding risk (score 4-7)	2 181 (11.7)	2194 (11.7)	.002

ACE = angiotensin-converting enzyme; ARB = angiotensin-receptor blocker; BMI=body mass index; $CHADS_2$ = stroke risk factor score based on Congestive heart failure, Hypertension, Age ≥ 75 years, Diabetes, prior Stroke; CPRD = Clinical Practice Research Datalink; IMD2015= Index of Multiple Deprivation 2015; IQR=interquartile range; LVEF=left ventricular ejection fraction; PSM = propensity score matching; SD=standard deviation; SE=systemic embolism; TIA=transient ischemic attack.

Table 5.3: Effectiveness Results in the CPRD Aurum ARISTOTLE-eligible matched cohort of new users (intent-to-treat)

All Patients	Apixaban Group (N=18 684)			Warfarin Group (N=18 684)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	439	32515	1.35	600	41003	1.46	0.91 (0.80, 1.03)
Stroke	397	32553	1.22	525	41075	1.28	0.94 (0.82, 1.07)
Ischemic or uncertain type of stroke	325	32594	1.00	378	41169	0.92	1.07 (0.92, 1.24)
Hemorrhagic stroke	84	32810	0.26	168	41363	0.41	0.63 (0.48, 0.81)
Systemic embolism	45	32817	0.14	81	41392	0.20	0.67 (0.47, 0.96)
Key secondary: death from any cause	1798	32856	5.47	2234	41466	5.39	1.01 (0.95, 1.08)
Other secondary outcomes							
Stroke, SE, or death from any cause	2115	32515	6.50	2613	41003	6.37	1.01 (0.96, 1.07)
Myocardial infarction	289	32605	0.89	332	41175	0.81	1.08 (0.92, 1.27)
Stroke, SE, myocardial infarction, or death from any cause	2321	32269	7.19	2816	40723	6.92	1.03 (0.98, 1.09)
Pulmonary embolism or deep-vein thrombosis	80	32795	0.24	147	41342	0.36	0.67 (0.51, 0.88)

Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, 2.5 years after the index date).

Table 5.4: Safety results in the CPRD Aurum ARISTOTLE-eligible matched cohort of new users

All Patients	Apixaban Group (N=18 684)			Warfarin Group (N=18 684)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary safety outcome: major bleeding	810	30942	2.62	1082	36161	2.99	0.86 (0.78, 0.94)
Intracranial	129	31521	0.41	208	36987	0.56	0.73 (0.59, 0.91)
Other location	175	31439	0.56	262	36874	0.71	0.77 (0.64, 0.94)
Gastrointestinal	535	31149	1.72	650	36530	1.78	0.94 (0.84, 1.05)
Net clinical outcomes							
Stroke, SE, or major bleeding	1124	30700	3.66	1417	35957	3.94	0.91 (0.84, 0.99)
Stroke, SE, major bleeding, or death from any cause	2438	30700	7.94	2699	35957	7.51	1.04 (0.98, 1.10)

Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, derived date of last exposure to index treatment).

Table 5.5: Effectiveness Results in the CPRD Aurum ARISTOTLE-eligible matched cohort of new users (on-treatment)

	Apixaban Group (N=18 684)			Warfarin Group (N=18 684)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	421	32515	1.29	514	41003	1.25	1.00 (0.88, 1.14)
Stroke	380	32553	1.17	450	41075	1.10	1.04 (0.90, 1.19)
Ischemic or uncertain type of stroke	313	32594	0.96	320	41169	0.78	1.19 (1.02, 1.40)
Hemorrhagic stroke	78	32810	0.24	149	41363	0.36	0.65 (0.50, 0.86)
Systemic embolism	43	32817	0.13	67	41392	0.16	0.76 (0.52, 1.12)
Key secondary efficacy outcome: death from any cause	1726	32856	5.25	1908	41466	4.60	1.13 (1.06, 1.20)
Other secondary outcomes							
Stroke, systemic embolism, or death from any cause	2034	32515	6.26	2246	41003	5.48	1.12 (1.06, 1.19)
Myocardial infarction	285	32605	0.87	306	41175	0.74	1.15 (0.98, 1.36)
Stroke, systemic embolism, myocardial infarction, or death from any cause	2242	32269	6.95	2444	40723	6.00	1.14 (1.07, 1.20)
Pulmonary embolism or deep-vein thrombosis	77	32795	0.23	130	41342	0.31	0.72 (0.54, 0.95)

Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, 2.5 years after the index date, derived date of last exposure to index treatment).

The on-treatment results (Table 5.5) showed evidence of attrition bias in the warfarin arm as was observed in the full emulation.

5.3. Summary

This chapter presented the results for the emulation of the ARISTOTLE RCT using UK electronic healthcare records (CPRD Aurum linked to HES and ONS). The results showed that the pre-specified benchmarking criteria were met and found non-inferiority of apixaban vs warfarin; however, superiority of apixaban vs warfarin (as seen in ARISTOTLE) was not observed. The results observed showed a slightly lower benefit for apixaban over warfarin in the UK population compared with ARISTOTLE results which may be due to the lower proportion of Asian patients and higher TTR of warfarin users in the UK when compared with ARISTOTLE. It is probable that the TTR of the warfarin users in this CPRD cohort may

not be representative of, and is likely to be higher than, the TTR that would be seen in the absence of alternative OAC treatments (before DOACs were available).

The successful emulation of ARISTOTLE suggests that the methods we applied to CPRD Aurum data are able to identify valid effect estimates for OACs used to treat AF. As a consequence, this increases our confidence in applying the same methods in excluded or under-represented groups treated with OACs for AF, and it is also possible they could be adapted to emulate reference trials in other therapeutic areas.

The results paper was followed by additional results from sensitivity analyses that showed that applying a minimum exposure requirement gave results consistent with the primary analysis. The results of the primary analysis in the prior VKA exposure strata were consistent and showed no evidence of selection bias in the prevalent users. An alternative analysis assessing the results in the ARISTOTLE-eligible population (without matching to the baseline characteristics of the ARISTOTLE population and restricting to new users) gave results consistent with the full emulation for effectiveness but showed a significant reduction in the risk of major bleeding on apixaban vs warfarin whereas this result was borderline significant in the full emulation. This difference can be explained by the larger sample size of the eligible new-user analysis when compared with the full emulation providing more precision in the treatment effect estimates (narrower confidence intervals). Other potential reasons may include a slightly lower TTR when restricting to new users, and potentially greater benefit of apixaban over warfarin in older and female patients that comprised a larger proportion of this cohort when compared with the full trial emulation.

The next chapter will present the results of using the methodological framework to explore the effectiveness and safety of apixaban vs warfarin in CPRD Aurum in special patient

populations excluded from ARISTOTLE (patients at increased risk of bleeding) or under-represented in ARISTOTLE (patients aged ≥ 75 years).

Chapter 6 Objective 3: Extension to excluded or under-represented patient groups

This chapter will describe the extension of the analysis to patient groups of interest using the methodological template developed in the emulation of ARISTOTLE:

- Patients aged ≥ 75 years that were under-represented in ARISTOTLE
- Patients at increased bleeding risk that were excluded by the ARISTOTLE eligibility criteria

6.1. Introduction

In the results presented in Chapter 5 of the emulation of ARISTOTLE in CPRD Aurum, the numbers of users of apixaban in CPRD Aurum excluded by the different exclusion criteria was presented and is shown below in Table 6.1.

Some of the exclusion criteria have an insufficient sample size to permit assessment of treatment effects. Other exclusion criteria with larger sample sizes are comprised of a diverse range of conditions such as the severe comorbid condition which includes groups such as patients with cancer with short median survival, dementia, and severe mental health conditions. To account for the range of diverse conditions contributing to the severe comorbid criteria it would be more useful for clinicians and patients to estimate treatment benefits and harms in 'meaningful' distinct subsets of this group such as in people with cancer and separately in people with dementia. However, cancer itself is an umbrella term comprising a wide range of conditions with different risks and outcomes making it difficult to account for confounding in this group, further complicated by the relatively small sample sizes that would result from looking in these subsets after applying the other exclusion criteria. For some criteria such as severe renal insufficiency, analysis would be complicated

by the high probability of a switch from DOAC to warfarin in patients with worsening renal function reflecting common practice during the study period.

Table 6.1 Number of patients in CPRD Aurum prescribed apixaban excluded by ARISTOTLE exclusion criteria

	Apixaban users * (N=67 539)
Excluded due to any ARISTOTLE exclusion criteria	26 052
AF reversible causes	2 625
Moderate or severe mitral stenosis	585
Increased bleeding risk	8 463
Other condition req. chronic anticoagulation	3 154
Persistent uncontrolled hypertension	1 602
Active infective endocarditis	52
Concomitant aspirin > 165 mg/day	31
Concomitant aspirin + thienopyridine	421
Severe comorbid condition[a]	8 312
Alcohol or drug abuse	2 263
Recent ischemic stroke (within 7 days)	462
Severe renal insufficiency	2 969
Elevated ALT, AST, or Total Bilirubin[b]	1 485
Platelet ≤ 100,000/ mm^3	536
Hemoglobin < 9 g/dL	1 095
Pregnant or breastfeeding[c]	48

* Patients in CPRD Aurum prescribed apixaban Jan2013 to July2019 that met the minimum registration and ARISTOTLE inclusion criteria.
AF = atrial fibrillation; ALT = alanine transaminase; AST = aspartate transaminase; CPRD = Clinical Practice Research Datalink; SES = socioeconomic status; ULN = upper limit of normal.
a Severe comorbid condition with life expectancy <1 year or reasons making participation impractical; b ALT or AST > 2X ULN or Total Bilirubin ≥ 1.5X ULN; c Pregnant or breastfeeding within 3 years prior

6.2. Under-represented patient group – patients aged ≥ 75 years

6.2.1. Results from ARISTOTLE

Patients aged ≥ 75 years were under-represented in ARISTOTLE relative to the expected age-distribution of the target population of patients with NVAF at increased stroke risk; In ARISTOTLE 35.5% patients were aged ≥ 75 years in the apixaban arm, compared with 62% of patients with the indication in CPRD Aurum prescribed apixaban. Similarly, patients aged 80-89 years (2352 patients, 12.9%) and >90 years (84 patients, 0.5%) were under-represented

in ARISTOTLE compared with people in CPRD prescribed apixaban meeting the eligibility criteria (36.2% aged 80-89 years and 8.1% aged >90 years). ARISTOTLE (42)performed a subgroup analysis for the primary efficacy and safety outcomes by age group (<65 years, 65 to < 75 years, ≥ 75 years) which showed no significant interaction between treatment and age group (p=0.12 for stroke/SE, p=0.64 for major bleeding) though there was suggestion of a slight trend towards greater benefit for apixaban vs warfarin in the older age groups. A later study analysing ARISTOTLE outcomes by age (134) found the absolute benefits of apixaban appeared to be greater in the elderly likely due to their higher risk of intracranial bleeding (Figure 6.1). RCTs are often underpowered for looking at treatment effects in subgroups given the RCTs tend to be designed to be powered for analysis of the primary outcome in the full cohort of patients.

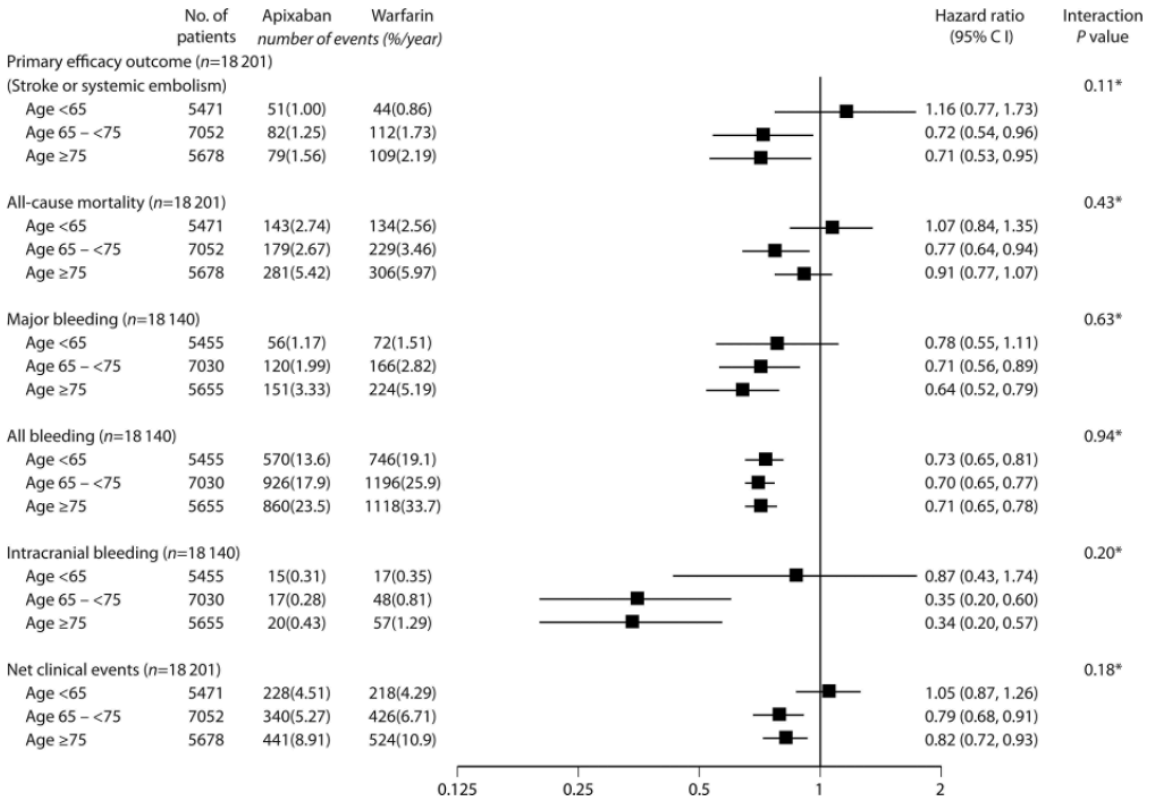


Figure 6.1 Effect of apixaban vs warfarin on major outcomes by age in ARISTOTLE

ref: Figure 2 The effect of apixaban vs. warfarin on major study outcomes according to age from 'Efficacy and safety of apixaban compared with warfarin according to age for stroke prevention in atrial fibrillation: observations from the ARISTOTLE trial.' Halvorsen S et al (134).
Eur Heart J, Volume 35, Issue 28, 21 July 2014, Pages 1864–1872, https://doi.org/10.1093/eurheartj/ehu046

The analysis of ARISTOTLE by age group suggested a slight trend towards greater benefit of apixaban vs warfarin in the elderly for bleeding outcomes however the underlying reasons for this trend is unclear and could relate to differences in TTR in the warfarin arm or reduced efficacy of reduced-dose of apixaban used in older patients meeting the criteria for dose-reduction; the large size of the cohort of elderly patients available in CPRD Aurum made it possible to see if a similar result would be observed in this cohort. There was an added benefit of being able to include a larger proportion of patients aged ≥80 years compared with the reference trial. AF prevalence is strongly related to age therefore determining the risks and benefits of different anticoagulants in the older cohort is important for clinicians and patients to know.

In the description of the epidemiology and treatment of AF in Chapter 2, age was identified as one of the most important risk factors for the development of AF and was also found to be associated with increased risks of stroke and bleeding events. Treatment decisions in elderly patients with AF can be difficult given the increased risk of both stroke and bleeding and the number of comorbidities in this group.

6.2.2. Methods for creation of the older age cohort in CPRD Aurum

The additional analysis presented in Chapter 5, in which ARSITOTLE-eligible new users of apixaban and warfarin were propensity score matched, included a large cohort of patients aged ≥ 75 years (22 548 patients, 60.3% of the patients) allowing the benefits and harms of apixaban vs warfarin in older people to be compared using this matched cohort.

6.2.3. Baseline comparison in the older age cohort

The baseline characteristics and balance between treatment groups of the older age group matched cohort and the younger cohort for comparison was assessed and is presented in

Table 6.2. Both age cohorts were well balanced across all characteristics assessed (mean standardised difference <0.1) indicating the propensity score matching had performed well at removing potential known sources of confounding.

Older patients were more likely to be female (52% vs 37%) and tended to have higher prevalence of all comorbidities other than hypertension and diabetes compared with the younger cohort. $CHADS_2$ stroke risk score was higher in the older patients (mean 2.6 vs 1.7) and the older patients were more likely to have renal impairment. Current smoking and alcohol consumption were higher in the younger cohort. Younger patients had higher use of antihypertensive medications.

Table 6.2 Baseline characteristics of the matched eligible new users of apixaban and warfarin in CPRD Aurum by age group

Characteristic	Younger age cohort (< 75 years)			Older age cohort (≥ 75 years)		
	Apixaban (N=7 406)	Warfarin (N=7 414)	Standardised difference	Apixaban (N=11 278)	Warfarin (N=11 270)	Standardised difference
Age - years, median (IQR)	68 (63, 71)	68 (63, 71)	.006	81 (78-85)	81 (78-85)	.009
Age <65 years	2 182 (29.6)	2 210 (29.8)	.005			
Age 65 to 74 years	5 214 (70.4)	5 204 (70.2)	.005			
Age 75 to 79 years				4 212 (37.3)	4 222 (37.5)	.002
Age 80 to 89 years				6 176 (54.8)	6 170 (54.7)	.000
Age ≥ 90 years				890 (7.9)	878 (7.8)	.004
Female sex-no.(%)	2 711 (36.6)	2 716 (36.6)	.001	5 818 (51.6)	5 770 (51.2)	.008
Systolic blood pressure - mm Hg, median (IQR)	130 (120, 40)	130 (120, 140)	.019	132 (121, 140)	132 (121, 140)	.012
Weight - kg, median (IQR)	90 (77, 104)	89 (77, 103)	.019	75 (65, 87)	76 (66, 87)	.031
Prior myocardial infarction - no. (%)	760 (10.3)	760 (10.3)	.000	1 399 (12.4)	1 400 (12.4)	.001
Prior clinically relevant or spontaneous bleeding – no.(%)	945 (12.8)	911 (12.3)	.014	1 881 (16.7)	1 880 (16.7)	.000
History of fall within previous year – no. (%)	24 (0.3)	35 (0.5)	.024	213 (1.9)	183 (1.6)	.020
Qualifying risk factors						
Age ≥ 75 years - no. (%)	0	0		11 278 (100.0)	11 270 (100.0)	
Prior stroke, TIA, or systemic embolism - no. (%)	1 353 (18.3)	1 410 (19.0)	.019	2471 (21.9%)	2448 (21.7%)	.005
Heart failure or reduced left ventricular ejection fraction - no. (%)	1 487 (20.1)	1 463 (19.7)	.009	2538 (22.5%)	2508 (22.3%)	.006
Diabetes - no. (%)	2 262 (30.5)	2 277 (30.7)	.004	2692 (23.9%)	2682 (23.8%)	.002
Hypertension requiring treatment - no. (%)	6 110 (82.5)	6 126 (82.6)	.003	8496 (75.3%)	8455 (75.0%)	.007
$CHADS_2$ score						
Mean	1.7 ± 0.9	1.7 ± 0.9	.008	2.6 ± 1.2	2.6 ± 1.2	.008
Distribution - no. (%)						
1	4 089 (55.2)	4 050 (54.6)	.012	1622 (14.4%)	1660 (14.7%)	.010
2	2 042 (27.6)	2 067 (27.9)	.007	4515 (40.0%)	4496 (39.9%)	.003
≥3	1 225 (16.5)	1 245 (16.8)	.007	5141 (45.6%)	5114 (45.4%)	.004

Characteristic	Younger age cohort (< 75 years)			Older age cohort (≥ 75 years)		
	Apixaban (N=7 406)	Warfarin (N=7 414)	Standardised difference	Apixaban (N=11 278)	Warfarin (N=11 270)	Standardised difference
3	866 (11.7)	883 (11.9)	.007	2 633 (23.3)	2 639 (23.4)	.002
4	285 (3.8)	294 (4.0)	.006	1 704 (15.1)	1 681 (14.9)	.005
5	74 (1.0)	68 (0.9)	.008	658 (5.8)	652 (5.8)	.002
6	0	0		146 (1.3)	142 (1.3)	.003
Medications at index date - no. (%)						
ACE inhibitor or ARB	4 536 (61.2)	4 509 (60.8)	.009	5 625 (49.9)	5 749 (51.0)	.023
Amiodarone	173 (2.3)	184 (2.5)	.010	170 (1.5)	170 (1.5)	.000
Beta-blocker	4 940 (66.7)	4 944 (66.7)	.000	6 551 (58.1)	6 494 (57.6)	.009
Aspirin	639 (8.6)	762 (10.3)	.056	1 132 (10.0)	1 186 (10.5)	.016
Clopidogrel	262 (3.5)	265 (3.6)	.002	448 (4.0)	469 (4.2)	.010
Digoxin	586 (7.9)	579 (7.8)	.004	1 036 (9.2)	1 094 (9.7)	.018
Calcium blocker	2 750 (37.1)	2 662 (35.9)	.025	3 470 (30.8)	3 541 (31.4)	.014
Statin	4 321 (58.3)	4 378 (59.1)	.014	6 031 (53.5)	5 987 (53.1)	.007
Nonsteroidal antinflammatory agent	418 (5.6)	419 (5.7)	.000	840 (7.4)	805 (7.1)	.012
Gastric antacid drugs	146 (2.0)	162 (2.2)	.015	325 (2.9)	301 (2.7)	.013
Proton pump inhibitor	2 364 (31.9)	2 413 (32.5)	.013	4 015 (35.6)	3 958 (35.1)	.010
H2 receptor antagonist	161 (2.2)	219 (3.0)	.049	446 (4.0)	392 (3.5)	.025
Renal function, creatine clearance - no. (%)						
Normal, >80 ml/min	4 598 (62.1)	4 723 (63.7)	.034	1 700 (15.1)	1 619 (14.4)	.020
Mild impairment, >50 to 80 ml/min	2 399 (32.4)	2 332 (31.5)	.020	5 751 (51.0)	5 792 (51.4)	.008
Moderate impairment (>30 to 50 ml/min)	333 (4.5)	281 (3.8)	.035	3 374 (29.9)	3 413 (30.3)	.008
Severe impairment (le 30 ml/min)	9 (0.1)	18 (0.2)	.028	383 (3.4)	362 (3.2)	.010
Not reported	67 (0.9)	60 (0.8)	.010	70 (0.6)	84 (0.7)	.015
Other risk factors and covariates						
Peripheral artery disease - no. (%)	386 (5.2)	387 (5.2)	.000	706 (6.3)	726 (6.4)	.007
Aortic plaque - no. (%)	1 313 (17.7)	1 385 (18.7)	.025	2 247 (19.9)	2 213 (19.6)	.007
Smoking status - no. (%)						

Characteristic	Younger age cohort (< 75 years)			Older age cohort (≥ 75 years)		
	Apixaban (N=7 406)	Warfarin (N=7 414)	Standardised difference	Apixaban (N=11 278)	Warfarin (N=11 270)	Standardised difference
Non-smoker	2 575 (34.8)	2 487 (33.5)	.026	4 309 (38.2)	4 331 (38.4)	.005
Ex-smoker	4 062 (54.8)	4 151 (56.0)	.023	6 450 (57.2)	6 434 (57.1)	.002
Current smoker	769 (10.4)	776 (10.5)	.003	519 (4.6)	505 (4.5)	.006
Alcohol consumption - no. (%)						
Non-drinker	2 194 (29.6)	2 251 (30.4)	.016	4 461 (39.6)	4 366 (38.7)	.017
Light drinker, up to 14 units per week	3 634 (49.1)	3 570 (48.2)	.018	5 691 (50.5)	5 812 (51.6)	.022
Moderate drinker, 15 to 42 units per week	1 382 (18.7)	1 407 (19.0)	.008	1042 (9.2)	1 004 (8.9)	.012
Heavy drinker, more than 42 units per week	196 (2.6)	186 (2.5)	.009	84 (0.7)	88 (0.8)	.004
Socioeconomic status - no. (%)						
England IMD2015 quintile 1(least deprived)	1 764 (23.8)	1 735 (23.4)	.010	2 985 (26.5)	2 968 (26.3)	.003
England IMD2015 quintile 2	1 719 (23.2)	1 662 (22.4)	.019	2 679 (23.8)	2 710 (24.0)	.007
England IMD2015 quintile 3	1 468 (19.8)	1 452 (19.6)	.006	2 283 (20.2)	2 335 (20.7)	.012
England IMD2015 quintile 4	1 257 (17.0)	1 337 (18.0)	.028	1 864 (16.5)	1 827 (16.2)	.009
England IMD2015 quintile 5(most deprived)	1 198 (16.2)	1 228 (16.6)	.010	1 467 (13.0)	1 430 (12.7)	.010
Ethnicity - no. (%)						
White	7 053 (95.2)	7 050 (95.1)	.007	10 896 (96.6)	10 894 (96.7)	.003
Black	69 (0.9)	80 (1.1)	.015	113 (1.0)	100 (0.9)	.012
South Asian	186 (2.5)	199 (2.7)	.011	176 (1.6)	161 (1.4)	.011
East Asian	16 (0.2)	17 (0.2)	.003	13 (0.1)	14 (0.1)	.003
Mixed	19 (0.3)	20 (0.3)	.003	21 (0.2)	23 (0.2)	.004
Other	25 (0.3)	20 (0.3)	.012	17 (0.2)	24 (0.2)	.015
Unknown	37 (0.5)	28 (0.4)	.018	42 (0.4)	54 (0.5)	.016
Charlson comorbidity index components - no. (%)						
Chronic obstructive pulmonary disease	768 (10.4)	786 (10.6)	.008	1 460 (12.9)	1 410 (12.5)	.013
Connective tissue disease	335 (4.5)	304 (4.1)	.021	817 (7.2)	848 (7.5)	.011
Peptic ulcer	278 (3.8)	295 (4.0)	.012	699 (6.2)	684 (6.1)	.005
Liver disease	37 (0.5)	67 (0.9)	.048	58 (0.5)	36 (0.3)	.030
Hemiplegia	17 (0.2)	17 (0.2)	.000	27 (0.2)	19 (0.2)	.016

Characteristic	Younger age cohort (< 75 years)			Older age cohort (≥ 75 years)		
	Apixaban (N=7 406)	Warfarin (N=7 414)	Standardised difference	Apixaban (N=11 278)	Warfarin (N=11 270)	Standardised difference
Cancer	690 (9.3)	727 (9.8)	.017	1 815 (16.1)	1 769 (15.7)	.011
Haematological cancer	104 (1.4)	121 (1.6)	.019	260 (2.3)	243 (2.2)	.010
BMI - kg/m^2, median (IQR)	30 (27, 35)	30 (27, 35)	.012	27 (24, 31)	27 (24, 31)	.016
ORBIT score						
Low bleeding risk (score 0-2)	6 924 (93.5)	6 949 (93.7)	.010	6 920 (61.4)	6 971 (61.9)	.010
Medium bleeding risk (score 3)	321 (4.3)	319 (4.3)	.002	2 338 (20.7)	2 251 (20.0)	.019
High bleeding risk (score 4-7)	161 (2.2)	146 (2.0)	.014	2 020 (17.9)	2 048 (18.2)	.007

ACE = angiotensin-converting enzyme; ARB = angiotensin-receptor blocker; BMI=body mass index; $CHADS_2$ = stroke risk factor score based on Congestive heart failure, Hypertension, Age ≥ 75 years, Diabetes, prior Stroke; CPRD = Clinical Practice Research Datalink; IMD2015= Index of Multiple Deprivation 2015; IQR=interquartile range; LVEF=left ventricular ejection fraction; PSM = propensity score matching; SD=standard deviation; SE=systemic embolism; TIA=transient ischemic attack.

6.2.4. Safety and effectiveness in the older age cohort

Results of the analyses of effectiveness and safety outcomes by age cohort are shown in Table 6.3. The older age group had higher absolute event rates for all outcomes, with annual rates of stroke/SE approximately double the rate seen in the younger age group (0.89 and 0.94 vs 1.67 and 1.83 in the younger and older apixaban vs warfarin) and around 3 times the rate of death. There was no evidence of any significant differences between the elderly group and younger patients in the treatment effects.

In the older age cohort, similar risks for apixaban vs warfarin were seen for stroke/SE (HR 0.90 [95% CI 0.77, 1.03]) and death (HR 1.00 [95% CI 0.93, 1.07]). Older people on apixaban had a lower risk of haemorrhagic stroke (0.65 [0.48, 0.89]), systemic embolism (0.55 [0.34, 0.88]), and pulmonary embolism or deep vein thrombosis (0.71 [0.52, 0.98]) than older people on warfarin.

For the safety outcomes higher event rates were seen in the older age cohort when compared with the younger patients. Apixaban was superior to warfarin for major bleeding in both age groups with older patients having a 10% lower risk when on apixaban and younger patients a 27% lower risk compared with warfarin. There were no significant interactions between treatment and age group, however, there was a consistent trend of lower hazard ratio estimates in the younger when compared with the older age groups.

INR control was slightly higher in the younger age cohort (43.4% vs 37.7% TTR >= 0.75, median TTR 0.751 vs 0.726) suggesting that any reduced benefit of apixaban compared with warfarin for major bleeding in the older age group could be due to other reasons such as apixaban dosing or concomitant medications.

Table 6.3 Outcomes in the matched trial-eligible patients in CPRD Aurum by age group

	Age group						P value for interaction
	Age < 75 years			Age ≥ 75 years			
Outcome	Apx Rate %/yr	Warf rate %/yr	Hazard ratio (95% CI)	Apx Rate %/yr	Warf rate %/yr	Hazard ratio (95% CI)	
Effectiveness outcomes							
Primary: Stroke/SE	0.89	0.94	0.93 (0.74, 1.18)	1.67	1.83	0.90 (0.77, 1.03)	0.76
Stroke	0.74	0.82	0.91 (0.70, 1.17)	1.55	1.60	0.95 (0.82, 1.11)	0.77
Ischemic or uncertain type of stroke	0.59	0.54	1.06 (0.79, 1.43)	1.28	1.18	1.07 (0.90, 1.27)	0.95
Haemorrhagic stroke	0.16	0.30	0.56 (0.34, 0.92)	0.32	0.48	0.65 (0.48, 0.89)	0.63
Systemic embolism	0.14	0.14	0.95 (0.52, 1.74)	0.13	0.23	0.55 (0.34, 0.88)	0.16
Death from any cause	2.46	2.32	1.06 (0.92, 1.23)	7.54	7.54	1.00 (0.93, 1.07)	0.45
Other secondary							
Stroke, SE or death from any cause	3.17	3.11	1.02 (0.89, 1.16)	8.81	8.68	1.01 (0.95, 1.08)	0.91
Myocardial infarction	0.63	0.53	1.16 (0.86, 1.57)	1.06	1.00	1.05 (0.87, 1.27)	0.58
Stroke, SE, MI, or death from any cause	3.73	3.51	1.05 (0.93, 1.19)	9.59	9.32	1.02 (0.96, 1.09)	0.66
Pulmonary embolism or DVT	0.15	0.26	0.56 (0.33, 0.95)	0.31	0.42	0.71 (0.52, 0.98)	0.49
Safety outcomes							
Primary safety: major bleeding	1.81	2.31	0.77 (0.65, 0.91)	3.18	3.47	0.90 (0.81, 0.999)	0.14
Intracranial	0.19	0.36	0.55 (0.34, 0.88)	0.56	0.70	0.79 (0.62, 1.02)	0.20
Other location	0.46	0.69	0.66 (0.48, 0.91)	0.62	0.73	0.85 (0.67, 1.08)	0.21
Gastrointestinal	1.18	1.31	0.88 (0.71, 1.09)	2.09	2.11	0.96 (0.84, 1.11)	0.49
Net clinical outcomes							
Stroke, SE, or major bleeding	2.51	2.92	0.84 (0.73, 0.98)	4.46	4.66	0.94 (0.86, 1.03)	0.23
Stroke, SE, major bleeding, or death from any cause	4.44	4.45	0.98 (0.88, 1.10)	10.38	9.66	1.06 (0.99, 1.13)	0.27

Apx=apixaban; CI=confidence interval; DVT=deep vein thrombosis; MI=myocardial infarction; SE=systemic embolism; Warf=warfarin;
Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, 2.5 years after the index date) for effectiveness outcomes and at the earliest of (outcome event, death, transfer out of practice, last collection date, derived date of last exposure to index treatment) for safety outcomes.

6.2.5. Discussion on the older age cohort analysis

Analysis of the outcomes of older patients in CPRD Aurum showed results broadly consistent with the reference trial emulation with similar risks of stroke/SE and all-cause death in patients on apixaban compared with warfarin and an approximately 10% lower risk of major bleeding on apixaban. A trend of slightly lower benefit of apixaban vs warfarin for bleeding

outcome was observed in the older age group when compared with the younger age group however no significant interaction between age group and treatment was detected.

Extension of the analysis to a patient group under-represented in the reference trial was straightforward by using the cohort of propensity score matched eligible users created during the emulation of ARISTOTLE. This analysis was limited by being in the ARISTOTLE-eligible subset of patients meaning these results may not be generalisable to older patients that would not have met the eligibility criteria.

The successful benchmarking of the emulation of ARISTOTLE in this data source before applying the methodological framework to look at the elderly patient group increases the confidence in the results obtained in this patient group. This extension shows the potential for providing high quality evidence in a patient group of interest without having to perform an RCT in the group of interest. The results observed in the elderly group support the NICE guidance on choice of oral anticoagulants for patients with NVAF and provides reassurance to older people taking these medications with both apixaban and warfarin showing similar effectiveness. The increased risks of both stroke and bleeding in the elderly group along with the greater burden of comorbidities and number of concomitant medications in this group leads to uncertainty on the choice of oral anticoagulant in this group. The results of this analysis showing similar effectiveness of apixaban vs warfarin in the elderly and a lower risk of major bleeding on apixaban compared with warfarin can aid decision making on choice of treatment in this group.

6.3. Excluded patient group – increased bleeding risk

The exclusion criteria ‘increased bleeding risk’ removed a large number of apixaban users; this criterion was interpreted as excluding patients with any of the following:

- haematological conditions leading to increased bleeding risk (such as haemophilia),
- recent major bleeding,

- recent haematuria, gynaecological, or other type of bleeding considered clinically relevant
- aneurysm or arteriovenous malformation,
- gastrointestinal conditions putting the patient at greater risk of bleeding (such as GI ulcers)
- gastrointestinal or brain tumours (though a subset of these would be further excluded by the severe comorbid exclusion criteria)

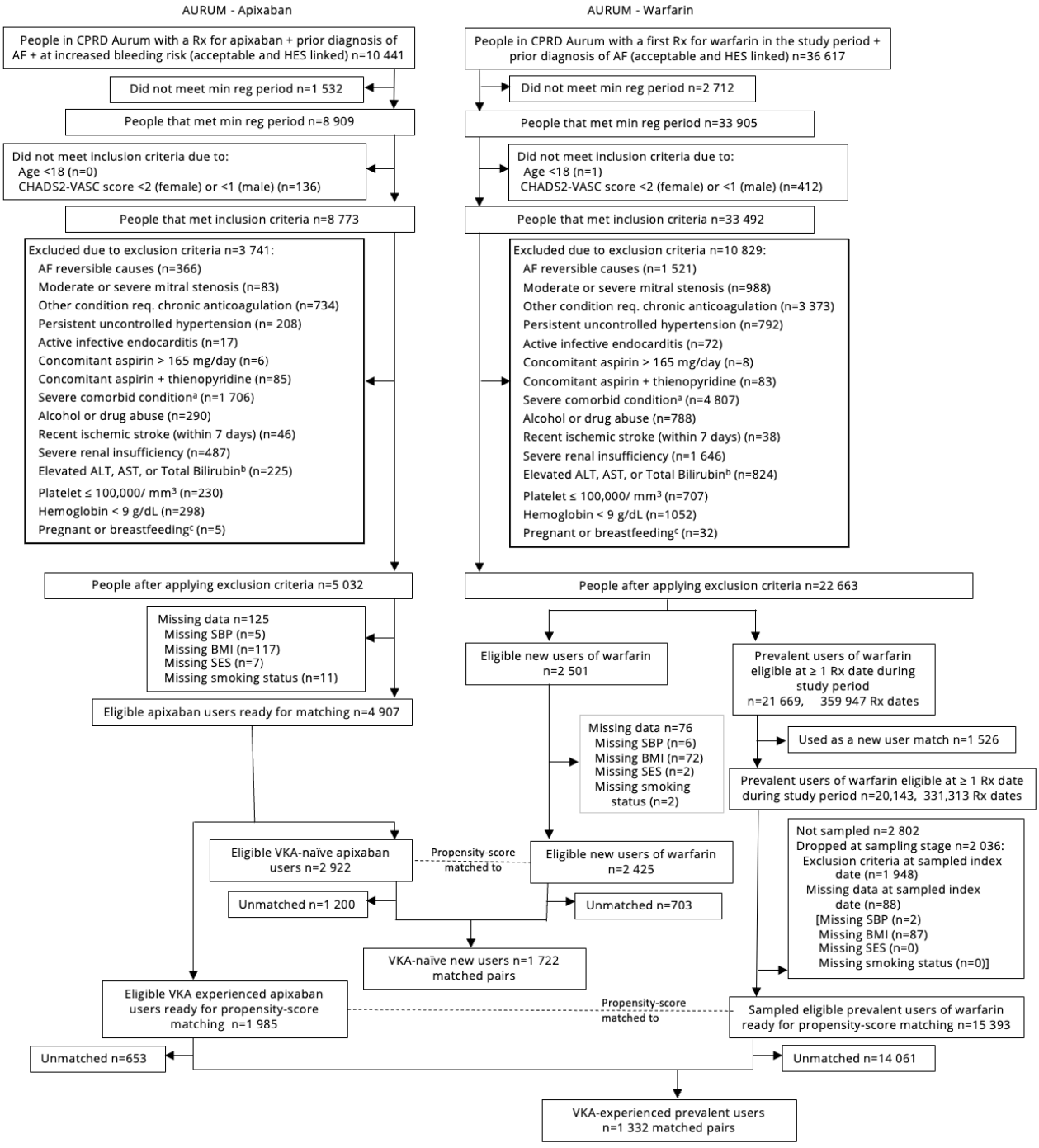
6.3.1. Methods for creation of the increased bleeding risk cohort in CPRD Aurum

An analysis was performed to estimate the effectiveness and safety of apixaban vs warfarin in patients with AF that were excluded by the increased bleeding risk criteria whilst also meeting other ARISTOTLE eligibility criteria. To increase the sample size and relevance to UK patients, the inclusion criteria from ARISTOTLE requiring at least one $CHADS_2$ stroke risk factor was replaced with the NICE guidance for when OACs are indicated in AF which uses the more updated stroke risk score (CHA_2DS_2VASc) requiring a CHA_2DS_2VASc score of 2 or above in women, and a CHA_2DS_2VASc score of 1 or above in men.

Restricting the cohort to new users of apixaban or warfarin only resulted in a small sample size (N=1 722 pairs) making it beneficial to include prevalent users to look at this patient group. The framework developed for the ARISTOTLE emulation was adapted by adding variables for the presence of the different increased bleeding risk factors to the propensity score models. Matching within the prevalent user strata was also simplified by using one propensity score model (not splitting by treatment history strata as this did not improve the balance or number of matched pairs) whilst maintaining the requirement for an exact match on categorised prior VKA exposure.

Figure 6.2 shows the selection of patients for this group. There were 8 773 patients prescribed apixaban and 33 492 patients prescribed warfarin that met the inclusion criteria of having AF, at increased bleeding risk, meeting the minimum registration period, being aged 18 years or older, and meeting the NICE criteria for OAC therapy for stroke prevention in AF. After applying the exclusion criteria there were 5 032 patients prescribed apixaban and 22 663 patients prescribed warfarin; propensity score matching resulted in 3 054 matched pairs. There were 1853 people (1200 new users and 653 switching from VKA to apixaban) at increased bleeding risk prescribed apixaban and eligible for matching for whom no match was found. The unmatched people exposed to apixaban (Table A4.1 in the appendix) tended to be older than those for whom a match was found (70.8% aged ≥ 75 years vs 65.5% of matched apixaban users), were more likely to be female (41.9% female vs 36.4% in matched cohort), more likely to have comorbidities (for example 19.6% vs 16.3% with COPD, 9.8% vs 6.6% with connective tissue disease) and more likely to have a haematological disorder as an increased bleeding risk factor (15.8% vs 6.8% of apixaban users that were matched).

Figure 6.2 Selection of cohort of patients at increased bleeding risk



Flow of number of individuals included in the analysis. AF = atrial fibrillation; ALT = alanine transaminase; AST = aspartate transaminase; BMI = body mass index; BP = blood pressure; CPRD = Clinical Practice Research Datalink; HES: Hospital Episodes Statistics; Rx = Prescription; SES = socioeconomic status; ULN = upper limit of normal; VKA = vitamin K antagonist.
a Severe comorbid condition with life expectancy <1 year or reasons making participation impractical; b ALT or AST > 2X ULN or Total Bilirubin ≥ 1.5X ULN; c Pregnant or breastfeeding within 3 years prior

Table 6.4 Baseline Characteristics of the Increased Bleeding Risk Cohort

Characteristic	Apixaban (N=3054)	Warfarin (N=3054)	Standardised difference
Age - years, median (IQR)	78 (72-84)	78 (72-84)	0.002
Female sex-no.(%)	1113 (36.4)	1100 (36.0)	0.009
Systolic blood pressure - mm Hg, median (IQR)	130 (120, 140)	130 (120, 140)	0.012
Weight - kg, median (IQR)	80 (69, 93)	80 (70, 93)	0.000
Prior myocardial infarction - no. (%)	612 (20.0)	608 (19.9)	0.003
Prior clinically relevant or spontaneous bleeding – no.(%)	1185 (38.8)	1188 (38.9)	0.002
History of fall within previous year – no. (%)	85 (2.8)	82 (2.7)	0.006
Prior use of vitamin K antagonist for >30 consecutive days - no. (%)	1332 (43.6)	1332 (43.6)	0.000
Qualifying risk factors			
Age ≥ 75 years - no. (%)	2001 (65.5)	2011 (65.8)	0.007
Prior stroke, TIA, or systemic embolism - no. (%)	1023 (33.5)	1026 (33.6)	0.002
Heart failure or reduced left ventricular ejection fraction - no. (%)	1042 (34.1)	1044 (34.2)	0.001
Diabetes - no. (%)	926 (30.3)	944 (30.9)	0.013
Hypertension requiring treatment - no. (%)	2319 (75.9)	2318 (75.9)	0.001
$CHADS_2$ score			
Mean	2.6 ± 1.4	2.6 ± 1.4	0.006
Distribution - no. (%)			
0	142 (4.6)	143 (4.7)	0.002
1	563 (18.4)	543 (17.8)	0.017
2	802 (26.3)	807 (26.4)	0.004
≥3	1547 (50.7)	1561 (51.1)	0.009
Medications at index date - no. (%)			
ACE inhibitor or ARB	1648 (54.0)	1678 (54.9)	0.020
Amiodarone	87 (2.8)	85 (2.8)	0.004
Beta-blocker	1930 (63.2)	1926 (63.1)	0.003
Aspirin	286 (9.4)	319 (10.4)	0.036
Clopidogrel	139 (4.6)	149 (4.9)	0.015
Digoxin	463 (15.2)	459 (15.0)	0.004
Calcium blocker	868 (28.4)	895 (29.3)	0.020
Statin	1910 (62.5)	1907 (62.4)	0.002
Nonsteroidal antinflammatory agent	198 (6.5)	185 (6.1)	0.018
Gastric antacid drugs	86 (2.8)	89 (2.9)	0.006
Proton pump inhibitor	1296 (42.4)	1244 (40.7)	0.035
H2 receptor antagonist	171 (5.6)	140 (4.6)	0.046
Renal function, creatine clearance - no. (%)			
Normal, >80 ml/min	830 (27.2)	847 (27.7)	0.012
Mild impairment, >50 to 80 ml/min	1369 (44.8)	1342 (43.9)	0.018
Moderate impairment (>30 to 50 ml/min)	747 (24.5)	750 (24.6)	0.002
Severe impairment (le 30 ml/min)	96 (3.1)	103 (3.4)	0.013
Not reported	12 (0.4)	12 (0.4)	0.000
Other risk factors and covariates			
Peripheral artery disease - no. (%)	473 (15.5)	500 (16.4)	0.024
Aortic plaque - no. (%)	1077 (35.3)	1097 (35.9)	0.014
Smoking status - no. (%)			
Non-smoker	924 (30.3)	954 (31.2)	0.021
Ex-smoker	1893 (62.0)	1872 (61.3)	0.014
Current smoker	237 (7.8)	228 (7.5)	0.011
Alcohol consumption - no. (%)			
Non-drinker	1143 (37.4)	1156 (37.9)	0.009
Light drinker, up to 14 units per week	1461 (47.8)	1444 (47.3)	0.011
Moderate drinker, 15 to 42 units per week	378 (12.4)	376 (12.3)	0.002
Heavy drinker, more than 42 units per week	34 (1.1)	38 (1.2)	0.012
Socioeconomic status - no. (%)			

Characteristic	Apixaban (N=3054)	Warfarin (N=3054)	Standardised difference
England IMD2015 quintile 1(least deprived)	796 (26.1)	796 (26.1)	0.000
England IMD2015 quintile 2	699 (22.9)	716 (23.4)	0.013
England IMD2015 quintile 3	594 (19.4)	611 (20.0)	0.014
England IMD2015 quintile 4	522 (17.1)	520 (17.0)	0.002
England IMD2015 quintile 5(most deprived)	443 (14.5)	411 (13.5)	0.030
Ethnicity - no. (%)			
White	2948 (96.5)	2942 (96.3)	0.011
Black	28 (0.9)	34 (1.1)	0.020
South Asian	52 (1.7)	50 (1.6)	0.005
East Asian	6 (0.2)	4 (0.1)	0.016
Mixed	6 (0.2)	9 (0.3)	0.020
Other	8 (0.3)	9 (0.3)	0.006
Unknown	6 (0.2)	5 (0.2)	0.008
Charlson comorbidity index components - no. (%)			
Chronic obstructive pulmonary disease	499 (16.3)	474 (15.5)	0.022
Connective tissue disease	203 (6.6)	205 (6.7)	0.003
Peptic ulcer	299 (9.8)	289 (9.5)	0.011
Liver disease	28 (0.9)	20 (0.7)	0.030
Hemiplegia	12 (0.4)	8 (0.3)	0.023
Cancer	569 (18.6)	555 (18.2)	0.012
Haematological cancer	84 (2.8)	76 (2.5)	0.016
BMI - kg/m^2, median (IQR)	28 (25, 32)	28 (25, 31)	0.010
Increased bleeding risk factor			
Aneurysm or AVM	1341 (43.9)	1358 (44.5)	0.011
Haematuria	621 (20.3)	627 (20.5)	0.005
Gastrointestinal bleed	702 (23.0)	718 (23.5)	0.012
Haematological disorder	209 (6.8)	211 (6.9)	0.003
Gynaecological bleed	163 (5.3)	137 (4.5)	0.039
Prior intracranial haemorrhage	119 (3.9)	119 (3.9)	0.000
Ocular bleed	40 (1.3)	39 (1.3)	0.003
Gastrointestinal or brain tumour	38 (1.2)	33 (1.1)	0.015
Other prior bleed	209 (6.8)	223 (7.3)	0.018
ORBIT score, median (IQR)	3 (2, 4)	3 (2, 4)	0.021
Low risk (0-2)	1269 (41.6)	1301 (42.6)	0.021
Medium risk 3	773 (25.3)	770 (25.2)	0.002
High risk ≥4	1012 (33.1)	983 (32.2)	0.020

ACE = angiotensin-converting enzyme; ARB = angiotensin-receptor blocker; BMI=body mass index; $CHADS_2$ = stroke risk factor score based on Congestive heart failure, Hypertension, Age ≥ 75 years, Diabetes, prior Stroke; CPRD = Clinical Practice Research Datalink; IMD2015= Index of Multiple Deprivation 2015; IQR=interquartile range; LVEF=left ventricular ejection fraction; PSM = propensity score matching; SD=standard deviation; SE=systemic embolism; TIA=transient ischemic attack; VKA = vitamin K antagonist;

After propensity score matching the treatment groups were well balanced on all baseline characteristics including presence of bleeding risk factors (Table 6.4), with a maximum mean standardised difference of 0.046. The proportion of people with prior clinically relevant bleeding in the apixaban arm of the increased bleeding risk cohort (38.8%, Table 6.4) was higher than in the ARISTOTLE emulation apixaban arm (17.3% in Table 3 of Results Paper

2), as was the proportion of those using proton pump inhibitors in the apixaban arm (42.4% in the increased bleeding cohort vs 34.5%) (Tables 6.4 and Paper 2 Table 3).

6.3.2. Results in the increased bleeding risk cohort

The results about the effectiveness in the increased bleeding risk group are displayed in Table 6.5. Similar results to the ARISTOTLE emulation in CPRD Aurum were seen in the increased bleeding risk group with a similar risk for apixaban and warfarin for stroke/systemic embolism (HR [95% CI] 0.94 [0.73, 1.20] in the increased bleeding risk group vs 0.98 [0.82, 1.19] in the ARISTOTLE emulation) and all-cause death (1.05 [0.93, 1.18. in the increased bleeding risk group vs 1.03 [0.93, 1.14] in the ARISTOTLE emulation).

Table 6.5 Outcomes in the matched patients in CPRD Aurum excluded by the increased bleeding risk group criteria

Outcome	Apixaban (N=3 054) Event rate %/yr	Warfarin (N= 3 054) Event rate %/yr	Increased Bleeding Risk Hazard ratio (95% CI)	*ARISTOTLE emulation Hazard ratio (95% CI)*
Effectiveness outcomes				
Primary: Stroke/SE	2.28	2.31	0.94 (0.73, 1.20)	*0.98 (0.82,1.19)*
Stroke	1.95	1.84	1.02 (0.78, 1.33)	*0.99 (0.81,1.21)*
Ischemic or uncertain type of stroke	1.45	1.32	1.06 (0.77, 1.45)	*1.13 (0.90,1.41)*
Haemorrhagic stroke	0.57	0.55	1.00 (0.61, 1.66)	*0.67 (0.44,1.01)*
Systemic embolism	0.32	0.52	0.56 (0.31, 1.03)	*1.01 (0.61,1.66)*
Death from any cause	9.72	9.11	1.05 (0.93, 1.18)	*1.03 (0.93,1.14)*
Other secondary				
Stroke, SE or death from any cause	3.17	3.11	1.05 (0.94, 1.17)	*1.04 (0.95,1.14)*
Myocardial infarction	1.41	1.29	1.09 (0.79, 1.51)	*1.01 (0.80,1.28)*
Stroke, SE, MI, or death from any cause	12.26	11.39	1.05 (0.94, 1.17)	*1.04 (0.96,1.14)*
Pulmonary embolism or DVT	0.30	0.47	0.59 (0.32, 1.12)	*0.65 (0.45,0.94)*
Safety outcomes				
Primary safety: major bleeding	4.90	5.19	0.92 (0.77, 1.10)	*0.88 (0.77,1.00)*
Intracranial	0.93	0.88	1.01 (0.68, 1.51)	*0.71 (0.51,1.00)*
Other location	0.99	1.23	0.79 (0.55, 1.15)	*0.93 (0.70,1.22)*
Gastrointestinal	3.12	3.19	0.94 (0.76, 1.18)	*0.88 (0.74,1.04)*
Net clinical outcomes				
Stroke, SE, or major bleeding	6.40	6.65	0.93 (0.80, 1.09)	*0.95 (0.84,1.06)*
Stroke, SE, major bleeding, or death from any cause	13.95	12.87	1.05 (0.95, 1.17)	*1.04 (0.96,1.13)*

CI=confidence interval; DVT=deep vein thrombosis; MI=myocardial infarction; SE=systemic embolism.
Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection

date, 2.5 years after the index date) for effectiveness outcomes and at the earliest of (outcome event, death, transfer out of practice, last collection date, derived date of last exposure to index treatment) for safety outcomes.

Safety results are included in Table 6.5 and also showed a similar risk of major bleeding for apixaban and warfarin (0.92 [0.77, 1.10] in the increased bleeding risk group vs 0.88 [0.77, 1.00] in the ARISTOTLE emulation).

6.3.3. Discussion on the increased bleeding risk cohort analysis

Analysis of the outcomes in the patients excluded by the increased bleeding risk criteria showed results similar to those obtained in the emulation of ARISTOTLE with a similar risk of stroke/SE, all-cause death, and major bleeding in apixaban compared with warfarin. The results from this analysis showing similar benefits and harms for apixaban vs warfarin in provides evidence to people in this patient group and supports the NICE guidance recommending apixaban as a treatment and warfarin as an option if apixaban is contra-indicated or not tolerated. Performing this analysis after benchmarking the reference trial provides more confidence in the results using these methods and in this data source of UK EHRs. The potential for increased risk of morbidity or mortality associated with bleeding in this patient group meant there was uncertainty on whether similar benefits and harms of apixaban compared with warfarin would be seen in this group. The results of this analysis may therefore be helpful in providing evidence in people with these conditions and show similar harms and benefits appear to apply in this group compared to patients eligible for the trial. This analysis was limited by the small sample size and the generalisability is limited by the relatively high proportion of unmatched patients in this cohort.

6.4. Summary

This chapter presented the results of the extension of the analysis evaluating benefits and harms of apixaban compared to warfarin in under-represented and excluded patient groups. The benchmarking of the emulation of ARISTOTLE using this data source and methods increases confidence in the results in these groups of interest. In this extension, we found the

extension to an under-represented group (patients aged ≥ 75 years) was simple to implement and could take advantage of the ARISTOTLE-eligible matched cohort of new users. The study of a patient group excluded from the emulation of ARISTOTLE (increased bleeding risk) proved more difficult given the wider range of comorbidities in this group meaning a large proportion of patients were excluded by other eligibility criteria and the diverse range of conditions included in the definition of the exclusion criteria necessitating the addition of extra variables to the propensity score. Despite including prevalent users, the increased bleeding group analysis was limited by a small sample size and a relatively high proportion of unmatched people. In both patient groups of interest similar risks of outcomes of interest were seen with apixaban vs warfarin and the results were consistent with those seen in the emulation of ARISTOTLE.

Chapter 7 Discussion

This chapter will provide an overall summary of the findings of this thesis that was focused on the emulation of a reference trial using UK electronic healthcare records data:

- The rationale for this work.
- The methods used
- Key results from the emulation of ARISTOTLE in CPRD Aurum and a comparison against the benchmarking criteria, trial results, and other relevant non-interventional study results
- The extension of the analysis to look at results by TTR and in underrepresented and excluded patient groups
- The key strengths and limitations of the study
- Future directions for research in this area
- Other work to come out of this thesis
- Conclusion

7.1. Summary of research and main findings

7.1.1. Methods

A protocol was published (BMJ Open) as part of this thesis which specified benchmarking criteria based on the ARISTOTLE results. The lower hazard ratio observed in the EU subgroup of the reference trial and reviewer comments in NICE questioning whether apixaban would be superior to well-controlled warfarin (TTR > 0.75) led to a plan to compare the TTR in the UK ARISTOTLE-analogous cohort to the reference trial and explore the impact of TTR as part of the study.

As part of the emulation of ARISTOTLE different methods were explored: i) the ability to match to the baseline characteristics of the reference trial without individual patient data, ii) different methods for the inclusion of prevalent users including the prevalent new user design and the forward sampling method, and iii) the impact of matching to the reference trial as opposed to omitting this step.

7.1.2. Results

Earlier in this thesis a detailed protocol was presented detailing the planned analysis for the emulation of ARISTOTLE in UK EHR data, additional detail on the methods work undertaken to perform this emulation, the results of the emulation (Chapter 5), and extensions to look at understudied or excluded patient groups in Chapter 6.

This thesis was successful in the emulation of ARISTOTLE using CPRD Aurum data with the results under final review for publication in *PLOS Medicine*. The methods used resulted in a cohort of patients that matched the ARISTOTLE participants on important baseline characteristics and the results for the primary effectiveness outcome met the benchmarking criteria.

Whilst non-inferiority was demonstrated in the CPRD cohort, superiority of apixaban vs warfarin for the primary endpoint was not seen in contrast to ARISTOTLE; this was likely due to a lower proportion of Asian patients, superior INR control, sub-optimal dosing of apixaban in some patients, and differences in concomitant medication use in the UK compared to the reference trial population. Omitting the step of matching to the trial participants gave results consistent with the VKA-naïve subset of the full emulation.

Table 7.1 Key comparison of results of the emulation of ARISTOTLE

Outcome Study	Apx Event Rate %/yr	Warf Event Rate %/yr	Hazard Ratio (95% CI)
Stroke or systemic embolism			
ARISTOTLE RCT	1.27	1.60	0.79 (0.66, 0.95)
ARISTOTLE RCT EU Subgroup			0.92 (0.56, 1.52)
Emulation of ARISTOTLE in CPRD Aurum	1.27	1.29	0.98 (0.82, 1.19)
ARISTOTLE-eligible new users CPRD Aurum	1.35	1.46	0.91 (0.80, 1.03)
ARISTOTLE-eligible new users RCT-DUPLICATE			0.73 (0.67, 0.79)
Death from any cause			
ARISTOTLE RCT	3.52	3.94	0.89 (0.80,0.998)
Emulation of ARISTOTLE in CPRD Aurum	4.37	4.20	1.03 (0.93, 1.14)
ARISTOTLE-eligible in CPRD Aurum	5.47	5.39	1.01 (0.95, 1.08)
Major bleeding			
ARISTOTLE RCT	2.13	3.09	0.69 (0.60, 0.80)
Emulation of ARISTOTLE in CPRD Aurum	2.45	2.77	0.88 (0.77, 1.00)
ARISTOTLE-eligible in CPRD Aurum	2.62	2.99	0.86 (0.78, 0.94)

The estimate of TTR showed the quality of INR control appeared to be higher in the CPRD cohort compared with ARISTOTLE and may explain some of the difference in benefits and harms of apixaban compared with warfarin observed in the UK population compared with the reference trial. The analysis by TTR showed a greater benefit for apixaban over warfarin in the low TTR group for major bleeding whereas in the high TTR group a higher risk of death was observed on apixaban compared with warfarin.

The analysis was extended to look at a patient group under-represented in the reference trial (patients aged ≥ 75 years) and an excluded patient group (increased bleeding risk). These analyses demonstrated similar results to those seen in the reference trial emulation with similar risks of stroke/SE and death for apixaban and warfarin and a slightly lower risk of major bleeding on apixaban.

7.2. Comparison with existing research

The literature review of non-interventional studies comparing apixaban to warfarin found only one other study that aimed to emulate ARISTOTLE(20); this study differed from the emulation performed in this thesis in the data source (US claims data), not matching to the

reference trial on baseline characteristics, and not including patients with prior VKA exposure ie applying a new-user approach. This emulation of ARISTOTLE in US data found a larger benefit for apixaban vs warfarin(20) as did the majority of the other non-interventional studies (12 other studies (86, 92-99, 115, 120, 122), with confidence intervals not overlapping with the confidence interval for the hazard ratio from my emulation of ARISTOTLE out of 21 reporting the primary endpoint) whereas results using EHRs from Sweden (110) and UK(111), 2 countries in which studies have found high quality of warfarin therapy (129) (140), showed treatment estimates closer to our results. The comparison of warfarin event rates and TTR reported in ARISTOTLE with those seen in the Aurum cohort suggested the difference in warfarin control quality between the trial participants and the Aurum patients likely contributed to the weaker treatment benefit of apixaban vs warfarin observed in the emulation in CPRD Aurum. This difference was somewhat surprising given that one typically expects trial conduct to result in superior quality of a treatment as a result of the greater focus on treatment adherence and protocol mandated regular monitoring of patients. Several of the RCTs comparing DOACs to warfarin were criticised for the quality of warfarin therapy. The EMA and NICE reviews on ARISTOTLE mentioned the TTR observed in ARISTOTLE was likely to be lower than may be relevant to EU/UK patients. Previous studies have found the UK to have high quality warfarin control, for example a study by Cotte et al (129) found 65.4% of patients in the UK in their study had TTR > 0.70 compared with only 47.8%, 44.2%, and 46.1% in France, Germany and Italy.

We found evidence of relatively good quality of warfarin control in the UK with a mean TTR of 0.73 in the CPRD Aurum emulation of ARISTOTLE. Sweden has also been observed as having excellent warfarin control potentially linked to the existence of a national quality registry for anticoagulation in AF that involves web-based algorithmic dose adjustment, a study in this registry found a mean TTR of 0.76 (140). In contrast studies looking at TTR in

the US have tended to report lower TTR, for example a large study of 140,000 patients in the US found an overall mean TTR of 0.54 (141).

The observation that the countries reported as having better quality INR control have also been those that have observed similar results showing a less dramatic benefit of apixaban vs warfarin in non-interventional studies lends weight to the argument that this is the main driver for the differences seen between studies in US data and UK and Swedish data.

7.3. Strengths

By using novel methodologies and a robust framework I was able to emulate a reference trial using routinely collected observational data and benchmark the results before extending the analysis to look at an under-represented group and an excluded group.

Proof of concept for construction of an RCT-analogous cohort without the use of individual participant data from the reference trial

The successful selection of a subset of patients matching the reference trial on aggregate provides a demonstration of how publicly available summary data of the reference trial can be used to support trial emulation where researchers wish to match to the RCT baseline characteristics. Without access to the individual patient data, I developed a novel method to select a cohort of patients in CPRD Aurum that matched the reference trial on baseline characteristics. The method involved constructing simultaneous equations describing the numbers of patients with different combinations of characteristics such as age group, sex, combinations of stroke risk factors, and renal function, followed by finding potential solutions to the equations via numerical estimation, and random sampling of patients from these subgroups using the numbers from the solution. The process was iteratively improved by comparing a contender solution of sampled patients against the ARISTOTLE apixaban arm at baseline and making changes to improve the match. The successful selection of a

subset of trial-eligible patients on apixaban that looked similar to the ARISTOTLE participants at baseline proves this to be a method worth considering for future reference trial emulation work in which researchers wish to match the baseline characteristics of the RCT and where the individual patient data are unavailable.

Inclusion of prevalent users

A particular feature of the reference trial emulated in this thesis involved the inclusion of prevalent users; exploring the feasibility of this in the reference trial emulation setting led to the finding that with the right method and careful approach to the application of eligibility criteria this can be achieved without introducing selection bias. This finding may be of relevance to certain areas:

- In the context of an existing 'gold standard' treatment, rare diseases, or rare subtypes of patients, where a large proportion of patients will already be on the treatment making the trial emulation more feasible when pre-existing users of the standard treatment are included in the study.
- For answering questions on the impact of initiating an 'add-on' therapy alongside a pre-existing standard treatment vs staying on the standard treatment without add-on treatment.
- In indications that by definition require prior treatment exposure such as treatment resistant depression.
- For answering questions on the impact of a new treatment vs no treatment since a similar approach to sampling of 'potential index dates' could potentially be used when looking at a control group on no treatment, though there are additional complexities and risks for bias when comparing to no treatment and selection of potential index dates in this scenario.

- To be able to answer questions for clinicians and patients on switching treatment or continuing conditional on prior treatment experience.

Compared with the method described by Webster-Clarke (136), the methods used in the reference trial emulation in this thesis added an extra step in checking the trial eligibility at the sampling stage (as the method in the simulation study paper did not state how to deal with this requirement). This method proved to be relatively easy to implement and offers an intuitive way of looking at eligibility for prevalent users given that it maps easily to the typical RCT screening process – in screening for an RCT the patient will likely meet the minimal key inclusion criteria prior to screening (for example meeting the simple inclusion criteria of having the disease or indication of interest and being aged > 18) to have been identified by an investigator as suitable for a study but then may have only 1 chance to pass screening during which all other eligibility criteria are checked. Some RCTs may allow re-screening for participants that fail time-dependent criteria such as prohibited concomitant medication or out of range laboratory values. Further research is required to understand if and how equivalent 're-screening' could be accounted for in a reference trial emulation study including prevalent users without introducing selection bias.

Despite the potential benefits to including prevalent users, doing so has the potential to make the study design and methods more complex and time-consuming compared with only including new users and increases the risk of selection bias.

As part of emulating a reference trial this thesis considered different methods for the inclusion of prevalent users and therefore provides a real-world example of the suitability of these relatively new methods in this area. The initial attempt to use the Suissa prevalent new user (PNU) design found this to be time consuming and unwieldy with propensity score models that would not converge. It was difficult to mimic the randomisation into the

reference trial since ineligible participants should fail screening and so should not be in the pool of potential matches; keeping ineligible patients in the pool requires substantial changes to the propensity score model and increases the difficulty in achieving a balanced cohort. The PNU method may be more suitable to studies including prevalent users that are not focused on reference trial emulation.

Successful emulation of ARISTOTLE in UK non-interventional data

A key strength of this thesis was the successful emulation of ARISTOTLE in UK non-interventional data by using novel methods both to select patients matching the trial participants and to include prevalent users of warfarin using routinely available EHR data. This provides a proof of concept and framework that can be adapted for future reference trial emulation studies.

The emulation of ARISTOTLE in this data source adds to the body of evidence in this therapeutic area and aids understanding of the potential causes for differences in results both between non-interventional studies and ARISTOTLE and between non-interventional studies.

Many phase III RCTs are multi-regional enrolling participants from a range of different countries across multiple continents; however, the subgroups of participants from each country or region will often be too small to draw conclusions on the treatment effectiveness in the individual regions given the RCT is likely to be powered for the overall population rather than any subgroups. A key point of interest when assessing the treatment effectiveness obtained when emulating a reference trial using EHRs may therefore be the degree to which any differences observed may have been caused by differences between the country or countries in which the EHRs were collected vs the countries in which the reference trial was conducted. There may be differences in the standards of care impacting areas such as regular monitoring of patients, dose optimisation, quality of comparative treatments administered,

and in the case of subjective or patient reported outcomes, cultural differences in patient expectations of treatment benefits can be important.

Analysis of TTR

A further strength was the extension of the emulation study to compare the TTR in the UK ARISTOTLE-analogous cohort with the reference trial and explore the potential impact of TTR on results.

The exploratory analysis by TTR suggested the relative benefits and harms of apixaban vs warfarin depended on the TTR of warfarin for some outcomes, with patients on apixaban having a lower risk of major bleeding compared with patients on poorly controlled warfarin in contrast to patients on apixaban having a slightly higher risk of death compared with well controlled warfarin.

7.4. Limitations

Missing baseline data

There was missing data for certain key covariates (namely renal function, BMI, alcohol consumption, smoking status, and socioeconomic status). For some of these variables a complete case approach was taken given the low proportion with missing data. For others such as alcohol intake a pragmatic approach of including a 'missing data' category was taken. The assumptions underlying the methods taken may not have been valid and additional methods such as imputation of missing baseline covariates using chained equations would have helped characterise the potential impact of this. For the prevalent users, there was a high rate of missing data for prior INR control making it difficult to include a variable relating to prior INR control in the propensity score model.

Residual confounding

Whilst matching methods were used in an attempt to account for confounding it is possible that residual confounding remains relating to variables omitted from the propensity score model, variables not measured (unmeasured confounding), or misspecification of the propensity score model. The data source used in this thesis enabled derivation of a wide range of variables potentially associated with the outcomes of interest meaning I could adjust for many confounders via propensity score methods; adjusting for these variables may not have made much difference to the effect estimates if they weren't strong confounders.

Inability to match on calendar time

In the therapeutic area addressed in this thesis there was a limited time window in which both treatments were prescribed as the newer DOACs were rapidly adopted as the first-line treatment option. In the ARISTOTLE-analogous cohort this led to a problem of being unable to match on calendar time and a resultant unequal follow-up time between the two treatment groups. This problem will be likely to occur irrespective of therapeutic area whenever a new treatment is rapidly adopted. A potential solution is to use a historical control cohort instead, however this option may cause issues of availability of other treatments and standards of care changing over time. The use of a PNU-style design as opposed to restricting to new users alone helps mitigate this problem as it allows inclusion of more comparator-exposed patient time in the period of adoption of the new drug.

An additional consideration is that requiring a match on calendar time could have increased bias in this study - if apixaban rapidly became the favoured first-line treatment then patients initiating warfarin later in study period are likely to be less similar to patients initiating apixaban compared with those patients that initiated warfarin earlier in the study period when warfarin was still more commonly selected as a first-choice treatment.

Uncertainty in the classification of exposure

A limitation of the data is in the uncertainty of classification of exposure. CPRD Aurum provides only GP prescription data meaning it is not known whether a patient filled the prescription at the pharmacy or took the medication. Prescriptions from hospital doctors are missing leading to uncertainty in when a patient may have first taken a medication or in treatment gaps involving hospitalisation during the follow-up of the patient. A further complication in accurately determining exposure is the overlap of treatments when a patient switches from one OAC to another. In such cases one cannot be sure the exact date a patient ceased taking the older medicine and started the newer one; with a further uncertainty introduced where delays in updating a GP repeat prescription system may result in a longer period of time with two different OACs prescribed.

Misclassification of concomitant medications is also possible for similar reasons and the medications recorded as ongoing at the index date may not accurately reflect the treatments taken in cases where medications likely to interact with the OAC such as aspirin may be discontinued at the index date but with a delay to this being reflected in the prescription record. This study attempted to account for this by requiring a prescription after the index date for treatments such as aspirin likely to have been discontinued on initiation of OAC therapy. Information on over-the-counter use of important concomitant medications such as aspirin, antacids, and NSAIDs is also missing leading to uncertainty in the derivation of these covariates. Moves in the UK to increase the range of medicines available over the counter (such as oral contraceptives) and increasing use of private healthcare could further reduce accurate classification of exposure and concomitant medications in future studies. Most of the people included in this study would have been eligible for free prescriptions meaning exposure to the index treatment and concomitant medications of interest are likely to have been recorded in the EHRs.

Difficulty in ascertainment of adherence

Another limitation of this study was the difficulty in ascertaining adherence. With UK data there is a reliance on using the number of prescriptions or number of tablets prescribed compared against the number of days covered. In this study attempting to measure adherence in this way did not appear helpful in discriminating high vs low adherence with nearly all apixaban users being estimated as having high adherence. One may hypothesise that the automatic issuing of prescriptions may mask the actual underlying use. For warfarin, most patients in CPRD Aurum did not have their daily dose recorded in a systematic way making it not possible to estimate adherence from the number of prescribed tablets against the number of days estimated in the treatment period.

Attrition bias in the warfarin arm

A treatment with a frequently sampled measure associated with the outcomes of interest will potentially be susceptible to attrition bias. For warfarin, the regular measurement of INR and calculation of TTR means clinicians would likely selectively choose patients to switch or continue based on their INR control leading to attrition bias and ultimately difficulty in emulating a reference trial in which the treatment is blinded and the protocol discourages treatment switching. A patient reporting side effects or more minor clinical manifestations such as minor bleeding could also lead a clinician to switch the patient's treatment. The attrition bias in the warfarin arm likely led to bias in both the intent-to-treat and on-treatment analyses as patients doing badly on warfarin most likely to have an event would be those most likely to switch to a DOAC during follow-up (in the absence of contraindications to DOACs); in the intent-to-treat analyses these patients may have experienced a lower event rate on the treatment they switched to compared to if they had stayed on warfarin whereas in the on-treatment analyses these patients may be censored before experiencing an event. In both the intent-to-treat and on-treatment analyses this treatment switching of patients more

likely to experience an event on warfarin to alternative DOAC would have led to under-estimating the beneficial effect of apixaban vs warfarin.

Selection bias in the warfarin arm

A further bias, related to attrition in the warfarin arm, is the risk of selection bias in the warfarin arm. Given the availability of alternative OACs during the study period it is possible that the patients on warfarin represent those patients more likely to do well on warfarin, in particular one may expect a survivorship bias in the prevalent users. Thus the results may underestimate the benefits of apixaban vs warfarin compared with a hypothetical situation in which patients are equally likely to initiate or remain on warfarin regardless of how well they do on warfarin (as measured by INR control or symptoms such as bleeding). The benchmarking of the results indicates this bias is unlikely to be a major issue.

Analysis by a post-baseline measure (TTR)

The analysis by TTR is limited given that TTR is observed post-index date and may be considered a proxy measure of adherence to warfarin treatment. A more suitable comparison may therefore have been to identify similarly highly-adherent apixaban users and perform a subgroup analysis grouping the low TTR warfarin users with the low-adherent apixaban users and the high-adherence apixaban users with the high-TTR warfarin users. In the CPRD Aurum data it was difficult to accurately measure apixaban adherence using measures such as proportion of days covered by prescriptions with a very low variability seen between apixaban users. The use of IPTW aimed to minimise the impact of any confounding on the treatment estimates obtained in the analysis by TTR. A further limitation of the analysis by TTR is the high likelihood of miss-classification of TTR for the patients missing TTR given the relatively low accuracy of the model in predicting TTR. Analysis by a post-baseline measure is likely to cause selection bias which may not be fully removed by the measures

taken (use of predicted TTR for those in which TTR was missing and IPTW). Given the risk of selection bias the analysis by TTR should be considered exploratory and interpreted with caution.

Lack of power/sample size to study excluded or under-represented patient groups

It was not possible to look at some of the under-represented or excluded patient groups of interest due to low sample size. Ethnicity was an important patient factor of interest given the under-representation of people of Black ethnicity in ARISTOTLE (only 227 patients – 1.2% of the ARISTOTLE population) and previously documented increased risk of bleeding associated with warfarin therapy is Asian people (142, 143). It was not possible to look at under-represented ethnic groups using the CPRD Aurum data due to the majority of the people in the study period being of white ethnicity. The low number of people of Black and Asian ethnicity in the CPRD cohort likely reflects the demographics in the UK of the older age group in this disease area combined with a lower prevalence of AF in people of Asian ethnicity (144); this limitation is therefore not likely to apply to conditions more common in younger age groups or more prevalent in a wider range of ethnicities.

Most of the exclusion criteria selected a limited number of people in which it would not be feasible to obtain meaningful treatment estimates. Some of the criteria that excluded larger numbers represented a diverse range of conditions, such as the severe comorbid criteria which covered people with life-limiting diseases such as terminal cancer, dementia, and severe mental health conditions. Although I was able to perform analysis on the increased bleeding risk group, results were generally underpowered.

7.5. Future directions

7.5.1. Selection of patients matching an RCT without individual patient data in a different therapeutic area

It would be of interest to see if the method used in this project to select a cohort matching the baseline characteristics of the reference trial on aggregate could be adapted for use in emulation of a reference trial in a different therapeutic area.

The selection of patient characteristics to include in the construction of the simultaneous equations will depend on the disease area and are likely to include age, sex, and key categories of disease characteristics and treatment history at baseline. Typically, these variables will be summarised in the baseline characteristics table of the RCT with researchers able to find further breakdowns based on key subgroup analyses in which characteristics by sex, age group, or disease status may be provided.

To make the method more robust one should consider obtaining multiple different solutions and performing repeat sampling by iterating across a number of different random seed numbers. The resulting treatment estimates could then be compared and averaged increasing confidence in the findings. In the emulation of ARISTOTLE this option was explored but was not completed due to the time involved in such an approach combined with the observation that in the CPRD cohort a large proportion of patients would always be selected due to certain subgroups requiring up to 100% sampling (and thereby meaning there would be little difference between different ‘solutions’).

In this thesis I also looked at the treatment effects in a cohort of trial-eligible new users without matching to the baseline characteristics of the reference trial. Similar results were obtained in this analysis compared with the full emulation in common with other reference trial emulation studies which have looked at the impact of including matching to the trial(33).

Whilst the use of the matching step may not be necessary in this therapeutic area, eliminating as many possible differences between the reference trial and the emulation in non-interventional data aids understanding of any differences in results seen.

7.5.2. Presentation of CONSORT diagrams and baseline tables when including prevalent users

When patients may be considered for inclusion in a cohort at multiple different dates, the optimal presentation of the CONSORT diagram and table of baseline characteristics requires further consideration. In this thesis the flow diagrams considered eligibility at the patient level, conditional on the hypothetical situation in which a patient be selected at a date on which they are eligible. For the sampling stage the flow diagram also presented the number of patients sampled at a date at which they were not eligible. This presentation, whilst sufficient to understand the selection of patients, may not be the ideal format for this type of study design and alternatives should be explored. When summarising baseline characteristics prior to applying eligibility criteria and prior to matching the most suitable method to summarise patients with multiple dates per patient is unclear. Consensus and guidance on this topic would aid understanding of future studies including prevalent users and comparison between studies including prevalent users.

7.5.3. Methods for the classification of prior treatment history

A key challenge in the use of routinely collected data arises from the complex treatment patterns observed in real-world data. When trying to determine treatment periods there are still unanswered questions on best way to do this. Methods using machine learning for classification may be more successful than simplistic algorithms that assign allowable gaps between prescriptions. A particular problem with UK data is not being able to ascertain if prescriptions were filled by a patient. A better algorithm could use all available information such as including hospitalisations and prescriptions issued in secondary care.

A topic requiring further research in the use of routinely collected data is the classification or characterisation of a patient's treatment history. A patient may have a treatment history showing exposure to multiple different treatments, unexplained treatment gaps, and time periods with overlapping treatments. Even with a trivial example of only 1 comparator drug being available at the same time as the study drug of interest it could be difficult to determine prior treatment duration where there are gaps between prescriptions in which it may not be clear whether the gap represents a true gap in exposure or some other reason for the gap such as prescriptions issued in secondary care or stockpiling by the patient. In the scenario where there are multiple drugs available for the indication it becomes more difficult to determine how to match patients on their prior treatment history. As well as total duration of prior exposure to each drug a researcher must also consider the start and stop times relative to the index date, and the fact that the exposure to 1 type of drug may be split into multiple distinct treatment periods punctuated by exposure to different drugs.

Further research is needed on the question of how to classify or match patients with complex prior treatment patterns. A suitable approach may be classification via algorithms adapted to the therapeutic area. Given the large volume of data a data-driven machine learning approach may be more suitable.

7.5.4. Further exploration of the sampling methods for inclusion of prevalent users

Further work exploring the classification and matching of prior treatment patterns is warranted, especially in therapeutic areas in which there are multiple treatments available and in which patients may have complicated treatment histories switching between many different treatments. The methods of inclusion of prevalent users described by Suissa(137) and Webster-Clarke(136) do not provide much guidance on this matter with the examples focused on prior exposure to only one treatment. Applying too strict a requirement on

matching exact prior treatment history may lead to sample sizes that are too small and cohorts not representative of the full patient population. The variety of and complexity in treatment patterns is highly dependent on therapeutic area and will relate to the number of treatments available and the typical patient journey. A further complication is the reason for switching being highly likely to be associated with effectiveness and/or tolerability.

A comment by Dell'Aniello(145) recommended that the Webster-Clark sampling method(136) (a method used in this thesis) would better reflect the original PNU design if the sampling were implemented in chronological order of the switcher's index date rather than in order of prior treatment duration. Future analyses could investigate the impact of sampling in this order in contrast to ordering by increasing length of prior exposure.

During the implementation of the sampling of continuing users of warfarin, the decision was made that a patient should only have one chance to pass screening. Whilst many RCTs allow a participant only one chance to be screened there are some RCTs that allow re-screening. The impact of allowing re-screening on the sampling of prevalent users whilst avoiding selection bias in trial emulation should be investigated.

7.5.5. Selection of prior treatment history strata

As part of the Webster-Clarke sampling method(136) patients are propensity score matched in treatment duration strata thereby allowing the predictors of treatment to vary according to prior treatment duration. This may better reflect real world practice in which we may hypothesise that a patient or clinician making the decision of whether to stay on an existing older treatment or switch to a newer treatment will be different if they have a long history of exposure to the older treatment or only a short history. In the ARISTOTLE emulation, the most suitable number of prior history strata was explored by trying different strata and assessing the number of matched pairs and baseline balance with the different options. This is

a topic that requires further research to guide future researchers in how to select the strata in this step.

7.5.6. Inclusion of historical control in reference trial emulation

Consider including time prior to the authorisation of the new drug in the study period

The study period planned for the emulation of ARISTOTLE included index dates from 01 Jan 2013 to 31 July 2019 with follow-up extending to 31 Jan 2020. This period was chosen to coincide with the period when both apixaban and warfarin were available for the indication of interest in the UK. During the cohort selection and matching process, it became apparent that apixaban had rapidly replaced warfarin as a preferred oral anticoagulant treatment choice greatly increasing the problem of channeling bias and making it not feasible to match apixaban and warfarin users on calendar date or year.

Although the use of propensity score matching should have minimised the risk of channelling bias, it is plausible that the pool of patients on warfarin deemed ARISTOTLE-eligible in the study period may have been doing better on warfarin than the mix of patients that would have been available prior to the availability of alternative OACs. This has implications for the understanding of the treatment effect estimates and likely means they may be less similar to the reference trial than desired.

By planning for inclusion of a time period prior to the availability of the new treatment, researchers can examine the impact of this potential problem on the treatment effect estimates and include this time period as either the primary or a sensitivity analysis.

For the ARISTOTLE emulation study one approach could be to apply the same methods but with warfarin users selected from the time period 2008 to 2013. The earlier warfarin period cohort arm matched to the apixaban users could be compared with the warfarin users selected in the contemporaneous period to examine how the warfarin cohort has changed. In the

ARISTOTLE example we may expect the earlier time period ARISTOTLE-eligible warfarin users to have a lower time in therapeutic range and less attrition bias due to the less common use of other treatments in the time period (though other DOACs such as dabigatran were available before apixaban for AF). Nonetheless, inclusion of the earlier time period and associated analyses may help understanding of the magnitude of any potential selection bias. For indications that already have several different treatment options prior to the availability of a new treatment of interest this approach may be less relevant.

7.5.7. Application of methods to account for treatment switching during follow-up

In this therapeutic treatment switching during follow-up was a particular concern given that i) switching from a DOAC to VKA may be related to the development of a contra-indication to DOACs and ii) switching from a VKA to a DOAC would be more likely for patients doing badly on warfarin. A comparison of patients that switched treatment against those that did not switch identified lower TTR in the patients that switched from warfarin to a DOAC. In order to better understand treatment switching the EHRs of patients that switch should be searched for evidence of development of contra-indications to DOACs and evidence of poor INR control in warfarin users.

Inverse probability of treatment censoring

A valuable extension to this study would be to apply inverse probability of censoring weighting to estimate the treatment effect that would have been seen in the absence of treatment switching. This approach relies on the existence of a sufficient number of patients in the cohort that continued on their index treatment despite having similar baseline characteristics and trajectory to the patients that switched treatment; this may be unlikely in practice meaning this is an area that requires further research.

Estimand strategy

When designing trial emulation studies an estimand strategy, such as a 'treatment policy' or 'hypothetical' estimand strategy, should be specified in the protocol to plan for how to handle events such as treatment switching. The estimand strategy would describe the meaning of the planned treatment estimates and the differences between the different planned estimands. The majority of non-interventional studies identified in the literature review censored patients at the time of treatment switching; the analysis of treatment switching in the emulation of ARISTOTLE in this thesis identified a higher rate of treatment switching in the warfarin arm compared to the apixaban arm and an observation of lower TTR in patients switching from warfarin to alternative OAC compared with those that did not switch indicating an assumption of non-informative censoring is likely invalid.

Use of an on-treatment approach alone may not be sufficient to account for treatment switching that may be related to markers of suboptimal effectiveness or safety on the index treatment or switching relating to the development of a contra-indication to one of the treatments that is also associated with higher risk of outcomes. The on-treatment approach may lead to informative censoring and potentially biased treatment effect estimates. A simple intent-to-treat approach may also lead to biased treatment estimates if patients doing badly on their index treatment are more likely to switch to a treatment on which they experience a lower risk of outcomes. More robust methods to account for intercurrent events during follow-up in reference trial emulation studies should be explored mirroring the estimand strategy approaches used in RCTs.

7.5.8. Use of alternative methods to address confounding

High dimensional propensity scores

This thesis selected variables and codelists based on the reference trial eligibility criteria and

baseline table, previous literature in the therapeutic area, and clinician insight into the important variables. An alternative would be a data driven approach such as high dimensional propensity scores(146, 147) which has the potential to be less biased than researchers selecting the codelists and variables to include in the propensity score models. Tazare et al(148) explored the use of high-dimensional propensity scores (HDPS) in UK EHRs and found adding the additional covariates identified by the high-dimensional propensity score algorithm resulted in a shirt in the propensity score distributions when compared with use of a conventional pre-defined covariate list suggesting that the HDPS had captured additional variables predictive of treatment. In general, more data driven machine learning approaches could help avoid or mitigate the dangers of researcher choice in the conduct of non-interventional studies; this can be balanced against the disadvantage of such approaches being limited by not knowing what is missing from the data or other contextual information relating to the data.

Inverse probability of treatment weighting

This thesis used propensity score matching (PSM) to create a balanced cohort for each objective. An alternative method using propensity scores is inverse probability treatment weighting (IPTW) (19) in which all patients would have been kept in the cohort with the contribution of their outcome data weighted so that the cohort is balanced. IPTW is more efficient than PSM as it makes use of all the patient data instead of dropping the data of patients that are not matched.

Study designs involving the inclusion of prevalent users made it more challenging to use IPTW given the methods proposed in the literature were based on propensity score matching. Future work exploring the use of IPTW within a PNU framework would be valuable, in the sampling method used in this study the full set of sampled continuing users could be matched

to the switchers using IPTW, potentially stratified by treatment history strata and/or coarsened exact strata based on prior treatment pattern.

Coarsened exact matching

This thesis adapted the concepts of coarsened exact matching (CEM)(16) for the selection of a subset of apixaban users in CPRD Aurum matching the trial population in aggregate, and trialled CEM for matching between the warfarin and apixaban arms with comparison against the balance obtained using propensity score matching in the new users (Appendix 3). Future work exploring incorporation of coarsened exact matching in the context of inclusion of prevalent users, and the optimal selection and parameterisation of variables to define the subgroups in the context of trial emulation would be helpful.

Cardinality matching

It was evident when attempting to find a selection of patients matching the trial that an ideal approach would have been to allow an algorithm to use combinatorics to form all possible matched cohorts and select the cohort giving the best balance weighted in order of preference of the most important variables. This was not possible at the time of the construction of the cohort due to limitations of computing power and the time such an approach would take. In addition, the combination of this approach with the methods allowing inclusion of prevalent users would have been complex.

Recent advances in computing power have given rise to cardinality matching introduced by Zubizarreta et al in 2014(149) and further described by Visconti et al(150), an approach that can construct the largest possible cohort that is matched meeting prespecified balance criteria. This method could prove useful in the emulation of reference trials by allowing researchers to set the criteria corresponding to matching to both the reference trial baseline characteristics and matching between treatment groups in the non-interventional cohort. Cardinality may be

more suitable when looking at excluded or under-represented patient groups with small sample size; a study by Fortin et al(151) found cardinality matching improved covariate balance in smaller sample sizes with limited covariate overlap compared with PSM.

7.5.9. Emulation of reference trials in different therapeutic areas

In this thesis there was success in emulating a reference trial in the setting of oral anticoagulants for the prevention of stroke in AF patients. At the start of the thesis when assessing the feasibility of a range of RCTs it was noticeable just how few therapeutic areas were feasible to replicate with the literature showing the majority of reference trial emulation studies to date being in a small selection of therapeutic areas – namely respiratory conditions, diabetes, OACs in AF and other conditions, statins for prevention of cardiovascular outcomes, and antihypertensives. By contrast there are few reference trial emulation studies in therapeutic areas such as psychiatry, oncology, and neurology. The main obstacle to conducting reference trial emulation studies in these other therapeutic areas using UK EHR data is a lack of suitable outcome data, a lack of exposure data for drugs given outside primary care, and in some cases small sample sizes.

The restriction of the lack of secondary care prescriptions can be solved by linking of secondary care prescriptions to the existing UK EHRs such as CPRD Aurum, HES, and ONS greatly expanding the range of treatments that could be studied.

A key example of missing outcome data covers outcomes such as patients reported outcomes (PRO), clinician-judged severity of symptoms or disease severity, and repeatedly measured outcomes on a continuous scale such as blood pressure. In an RCT when such measures are used as the outcome measure they will tend to be recorded at a baseline timepoint and then at certain pre-specified timepoints post-treatment initiation; thus facilitating a comparison between treatments in the change from baseline to a certain timepoint in the outcome. Whilst

some of these outcome measures are recorded in UK EHRs, the timing and frequency of their recording does not tend to correspond to RCT outcome timings. Both the baseline and outcome measure may be absent in EHRs and any measurements that are present will tend to be recorded at non-regular timepoints when comparing between patients on the same drug. Furthermore, reflecting on when or why such an outcome measure is or is not recorded for a given patient leads to difficult questions of missing data and suitable assumptions regarding missing at random or missing not at random.

To expand the field of reference trial emulation to a wider range of therapeutic areas and subsequently enable studies of drug effectiveness in these therapeutic areas more broadly this issue of missing outcome data requires a solution. Potential solutions include improving recording of the target reference trial outcome measures in EHR (for example by incentivising the assessment of disease severity at certain timepoints), or use of alternative outcome measures such as time to event.

Another challenge noticed during the feasibility work of the thesis was the problem of poorly recorded disease characteristics for some conditions. This problem may be solved by directly linking outpatient secondary care data to the existing UK EHRs or by classification of the disease using all available patient data, for example by machine learning. A pragmatic solution to the problem of missing disease type where this is missing for only a relatively small proportion of patients could be restricting the patient cohort to those with disease type well recorded; however, this has the disadvantage that the resulting patient cohort may not be a representative sample of all patients with the condition.

7.6. Other work to come out of this thesis

Two additional studies were completed using the derived datasets created for this thesis:

i) A study investigating moderation of the effectiveness and safety of apixaban vs warfarin by obesity in patients with atrial fibrillation

ii) A study looking at whether DOACs are associated with a lower risk of incident diabetes compared with warfarin in patients with atrial fibrillation

7.6.1. Moderation of the effectiveness and safety of apixaban vs warfarin by obesity in patients with atrial fibrillation

This study looked at the moderation of the comparative effectiveness and safety of apixaban versus warfarin by obesity in non-valvular atrial fibrillation. The study question aimed to address concerns that the fixed-dosing of apixaban may lead to under-dosing of obese patients and resultant lower effectiveness of apixaban in these patients.

Bin Hammad emulated a target trial similar to ARISTOTLE and compared 36-month risk ratios for stroke/SE, major bleeding, and all-cause mortality between apixaban and warfarin across different BMI strata (normal weight, overweight, obese). No difference was observed in the effect of apixaban vs warfarin on the risk of stroke/SE across the BMI strata whereas for major bleeding apixaban appeared more effective than warfarin in the overweight and obese groups (though no significant interactions were observed and all confidence intervals overlapped).

7.6.2. Risk of incident diabetes in DOACs compared with warfarin in patients with atrial fibrillation

The second study that came out from this thesis was inspired by findings in Hong Kong(152) and Taiwan(153), both showing a reduced risk of type II diabetes (T2DM) in DOAC users compared with warfarin users. Several studies have shown an association between vitamin K levels in blood, insulin sensitivity, and measures of glucose control as summarised in a review by Ho et al 2020 (154) providing a plausible biological mechanism for the observed association. The Yan-Ling study using data from this thesis sought to see whether the

findings would be replicated in UK EHR data with a much larger sample size, and in addition to investigate whether there was effect modification by sex or age group. The Yan-Ling study included patients with AF that were new users of warfarin or DOAC in the study period and had no history of diabetes; patients were followed-up until first diagnosis of T2DM or censoring. A competing risks approach was used with treatment compared via Cox proportional hazards models adjusted for multiple confounders. Yan-Ling found a lower risk of incident T2DM in DOAC users vs warfarin [aHR 0.90 (0.84, 0.95)] consistent with the results from Taiwan and Hong Kong. There was no evidence of effect interaction by age group or sex, in contrast to the Taiwan study which found a benefit for DOAC in lower risk of incident diabetes only in patients aged 65 years and over.

7.7. Conclusions

The result of the emulation of ARISTOTLE in UK EHRs leads to a few conclusions: Firstly that UK EHRs can be used to successfully emulate a reference trial in the therapeutic area of oral anticoagulants for the prevention of stroke in NVAF. Secondly, that the applicability of results of an RCT to a population may be influenced by differences in the quality of the standard of care, ethnicity, and typical use of other medications, between the countries where the reference trial took place and the country of the population of interest. Furthermore, in therapeutic areas where there is significant variability in standard of care country-specific non-interventional studies are likely to be of most relevance to a regulator or payer. This thesis presented a novel method for the selection of patients matching the trial participants at baseline on aggregate without access to individual patient data that can be adapted in other reference trial emulation studies. A further conclusion from the success of the emulation of ARISTOTLE is that the inclusion of prevalent users in trial emulation studies is feasible and can be implemented whilst avoiding the introduction of selection bias

using the Webster-Clarke sampling method with trial eligibility assessed at the point of sampling.

There is an ethical argument for requiring the study of benefits and harms of treatments in under-represented or excluded patient groups. The complexity of the methods involved in conducting trial emulation studies, (a combination of multiple techniques may be required to properly account for all potential sources of bias, confounding by indication, and treatment switching during follow-up), poses a risk of bias should an investigator have a vested interested in a particular finding. Whilst publishing a protocol and analysis plan prior to conducting a trial emulation study can help to limit this risk, it can be difficult to anticipate all possible data quality issues and sources of bias that may not become apparent until after starting an analysis. As the number of reference trial emulation studies published increases these limitations should reduce with researchers able to use the methods, findings, and limitations of prior reference trial emulations to guide the optimal design of future studies.

References

1. Powell EM, Douglas IJ, Gungabissoon U, Smeeth L, Wing K. Real-world effects of medications for stroke prevention in atrial fibrillation: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results. BMJ Open. 2021;11(4):e042947.
2. Ha C, Ullman TA, Siegel CA, Kornbluth A. Patients enrolled in randomized controlled trials do not represent the inflammatory bowel disease patient population. Clin Gastroenterol Hepatol. 2012;10(9):1002-7; quiz e78.
3. Foldesi C, Penzes M, Szabo M. Comparing Characteristics of Hungarian "Real Life" Patients With Participants of RCTS: First-Line Disease Modifying Therapies in Multiple Sclerosis. Value Health. 2014;17(7):A404.
4. Diao D, Wright JM, Cundiff DK, Gueyffier F. Pharmacotherapy for mild hypertension. Cochrane Database Syst Rev. 2012;2012(8):Cd006742.
5. Agency EM. Scientific guidance on post-authorisation efficacy studies. 2017.
6. Agency EM. Post-authorisation safety studies (PASS). 2020.
7. Sendor R, Stürmer T. Core concepts in pharmacoepidemiology: Confounding by indication and the role of active comparators. Pharmacoepidemiol Drug Saf. 2022;31(3):261-9.
8. Ahn N, Nolde M, Günter A, Güntner F, Gerlach R, Tauscher M, et al. Emulating a target trial of proton pump inhibitors and dementia risk using claims data. Eur J Neurol. 2022;29(5):1335-43.
9. Delapaz NR, Hor WK, Gilbert M, La AD, Liang F, Fan P, et al. An Emulation of Randomized Trials of Administrating Antipsychotics in PTSD Patients for Outcomes of Suicide-Related Events. J Pers Med. 2021;11(3).
10. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.
11. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol. 2016;79:70-5.
12. Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. Int J Epidemiol. 1986;15(3):413-9.
13. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999;10(1):37-48.
14. Lobo FS, Wagner S, Gross CR, Schommer JC. Addressing the issue of channeling bias in observational studies with propensity scores analysis. Research in Social and Administrative Pharmacy. 2006;2(1):143-51.
15. Williamson EJ, Forbes A. Introduction to propensity scores. Respirology. 2014;19(5):625-35.
16. Iacus SM, King G, Porro G, . Causal Inference without Balance Checking: Coarsened Exact Matching. Cambridge University Press; 2012. p. 1-24.
17. Rosenbaum PR, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. Journal of the Royal Statistical Society: Series B (Methodological). 1983;45(2):212-8.
18. Williamson E, Morley R, Lucas A, Carpenter J. Propensity scores: from naive enthusiasm to intuitive understanding. Stat Methods Med Res. 2012;21(3):273-93.
19. Rosenbaum PR. Model-Based Direct Adjustment. Journal of the American Statistical Association. 1987;82(398):387-94.

20. Wang SV, Schneeweiss S, Franklin JM, Desai RJ, Feldman W, Garry EM, et al. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. Jama. 2023;329(16):1376-85.
21. Baptiste P. Real-world effectiveness and adverse events caused by ACE inhibitors and ARBs

for reduction in cardiovascular events with validation against the ONTARGET trial. London School of Hygiene & Tropical Medicine: London School of Hygiene & Tropical Medicine; 2023.
22. Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT DUPLICATE Initiative. Circulation. 2021;143(10):1002-13.
23. Fralick M, Kesselheim AS, Avorn J, Schneeweiss S. Use of health care databases to support supplemental indications of approved medications. 2018;178:55-63.
24. Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L. Real world effects of COPD medications: a cohort study with validation against RCT results. 2021;57:2001586.
25. Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, et al. Medications for chronic obstructive pulmonary disease: a historical non-interventional cohort study with validation against RCT results. 2021;25:51.
26. Matthews AA, Szummer K, Dahabreh IJ, Lindahl B, Erlinge D, Feychting M, et al. Comparing Effect Estimates in Randomized Trials and Observational Studies From the Same Population: An Application to Percutaneous Coronary Intervention. J Am Heart Assoc. 2021;10(11):e020357.
27. Matthews AA, Dahabreh IJ, Fröbert O, Lindahl B, James S, Feychting M, et al. Benchmarking Observational Analyses Before Using Them to Address Questions Trials Do Not Answer: An Application to Coronary Thrombus Aspiration. American Journal of Epidemiology. 2022;191(9):1652-65.
28. Yiu ZZN, Mason KJ, Hampton PJ, Reynolds NJ, Smith CH, Lunt M, et al. Randomized Trial Replication Using Observational Data for Comparative Effectiveness of Secukinumab and Ustekinumab in Psoriasis: A Study From the British Association of Dermatologists Biologics and Immunomodulators Register. JAMA Dermatology. 2021;157(1):66-73.
29. Merola D, Young J, Schrag D, Lin KJ, Robert N, Schneeweiss S. Oncology Drug Effectiveness from Electronic Health Record Data Calibrated Against RCT Evidence: The PARSIFAL Trial Emulation. Clin Epidemiol. 2022;14:1135-44.
30. Boyne DJ, Cheung WY, Hilsden RJ, Sajobi TT, Batra A, Friedenreich CM, et al. Association of a Shortened Duration of Adjuvant Chemotherapy With Overall Survival Among Individuals With Stage III Colon Cancer. JAMA Netw Open. 2021;4(3):e213587.
31. Rizvi SL, Hughes CD, Hittman AD, Vieira Oliveira P. Can Trainees Effectively Deliver Dialectical Behavior Therapy for Individuals With Borderline Personality Disorder? Outcomes From a Training Clinic. J Clin Psychol. 2017;73(12):1599-611.
32. Baptiste PJ, Wong AY, Schultze A, Clase CM, Leyrat C, Williamson E, et al. Cardiorenal Effects of Angiotensin-converting enzyme inhibitors and Angiotensin receptor blockers in people underrepresented in trials: analysis of routinely collected data with validation against a target trial. medRxiv. 2022:2022.12.02.22282220.
33. Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, et al. Real world effects of COPD medications: a cohort study with validation against results from randomised controlled trials. Eur Respir J. 2021;57(3).

34. Heyard R, Held L, Schneeweiss S, Wang SV. Design differences and variation in results between randomised trials and non-randomised emulations: meta-analysis of RCT-DUPLICATE data. BMJ Med. 2024;3(1):e000709.
35. Held L, Micheloud C, Pawel S. The assessment of replication success based on relative effect size. The Annals of Applied Statistics. 2022;16(2):706-20, 15.
36. Baptiste PJ, Wong AY, Schultze A, Clase CM, Leyrat C, Williamson E, et al. Comparative effectiveness of ARB and ACEi for cardiovascular outcomes and risk of angioedema among different ethnic groups in England: an analysis in the UK Clinical Practice Research Datalink with emulation of a reference trial (ONTARGET). medRxiv. 2024:2024.01.17.24301397.
37. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. 2019;48:1740-g.
38. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). 2015;44:827-36.
39. Carbonari DM, Saine ME, Newcomb CW, Blak B, Roy JA, Haynes K, et al. Use of demographic and pharmacy data to identify patients included within both the Clinical Practice Research Datalink (CPRD) and The Health Improvement Network (THIN). Pharmacoepidemiology and Drug Safety. 2015;24(9):999-1003.
40. (CSDR) Cc. [Available from: https://clinicalstudydatarequest.com/.
41. NICE. Apixaban for preventing stroke and systemic embolism in people with non-valvular atrial fibrillation. Technology appraisal guidance. Reference number:TA275. Published: 27 February 2013 Last updated: 02 July 2021. 2013.
42. Granger CB, Alexander JH, McMurray JJV, Lopes RD, Hylek EM, Hanna M, et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. New England Journal of Medicine. 2011;365:981-92.
43. Golwala H, Jackson LR, Simon DN, Piccini JP, Gersh B, Go AS, et al. Racial/ethnic differences in atrial fibrillation symptoms, treatment patterns, and outcomes: Insights from Outcomes Registry for Better Informed Treatment for Atrial Fibrillation Registry. American Heart Journal. 2016;174:29-36.
44. Gahungu N, Trueick R, Coopes M, Gabbay E. Paroxysmal atrial fibrillation. BMJ. 2021;375:e058568.
45. NICE. Atrial fibrillation clinical knowledge summary 2023.
46. Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, et al. Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. Circulation. 2004;110(9):1042-6.
47. Schnabel RB, Yin X, Gona P, Larson MG, Beiser AS, McManus DD, et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. The Lancet. 2015;386(9989):154-62.
48. Psaty BM, Manolio TA, Kuller LH, Kronmal RA, Cushman M, Fried LP, et al. Incidence of and risk factors for atrial fibrillation in older adults. Circulation. 1997;96(7):2455-61.
49. Cadby G, McArdle N, Briffa T, Hillman DR, Simpson L, Knuiman M, et al. Severity of OSA is an independent predictor of incident atrial fibrillation hospitalization in a large sleep-clinic cohort. Chest. 2015;148(4):945-52.
50. Buch P, Friberg J, Scharling H, Lange P, Prescott E. Reduced lung function and risk of atrial fibrillation in the Copenhagen City Heart Study. Eur Respir J. 2003;21(6):1012-6.
51. Baber U, Howard VJ, Halperin JL, Soliman EZ, Zhang X, McClellan W, et al. Association of chronic kidney disease with atrial fibrillation among adults in the United States: REasons for Geographic and Racial Differences in Stroke (REGARDS) Study. Circ Arrhythm Electrophysiol. 2011;4(1):26-32.

52. Xiang H, Xue Y, Chen Z, Yu Y, Peng Y, Wang J, et al. The Association Between Left Ventricular Hypertrophy and the Occurrence and Prognosis of Atrial Fibrillation: A Meta-Analysis. Frontiers in Cardiovascular Medicine. 2021;8.
53. Okin PM, Wachtell K, Devereux RB, Harris KE, Jern S, Kjeldsen SE, et al. Regression of Electrocardiographic Left Ventricular Hypertrophy and Decreased Incidence of New-Onset Atrial Fibrillation in Patients With Hypertension. JAMA. 2006;296(10):1242-8.
54. Chung H, Lee JM. Left Atrial Remodeling and Thromboembolic Risk in Patients With Atrial Fibrillation. Int J Heart Fail. 2022;4(1):26-8.
55. Samokhvalov AV, Irving HM, Rehm J. Alcohol consumption as a risk factor for atrial fibrillation: a systematic review and meta-analysis. Eur J Cardiovasc Prev Rehabil. 2010;17(6):706-12.
56. Watson T, Shantsila E, Lip GYH. Mechanisms of thrombogenesis in atrial fibrillation: Virchow's triad revisited. The Lancet. 2009;373(9658):155-66.
57. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. 1991;22(8):983-8.
58. Friberg L, Rosenqvist M, Lindgren A, Terént A, Norrving B, Asplund K. High prevalence of atrial fibrillation among patients with ischemic stroke. Stroke. 2014;45(9):2599-605.
59. Ali AN, Abdelhafiz A. Clinical and Economic Implications of AF Related Stroke. J Atr Fibrillation. 2016;8(5):1279.
60. Andrew NE, Thrift AG, Cadilhac DA. The prevalence, impact and economic implications of atrial fibrillation in stroke: what progress has been made? Neuroepidemiology. 2013;40(4):227-39.
61. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of Clinical Classification Schemes for Predicting StrokeResults From the National Registry of Atrial Fibrillation. JAMA. 2001;285(22):2864-70.
62. 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;137(2):263-72.
63. O'Brien EC, Simon DN, Thomas LE, Hylek EM, Gersh BJ, Ansell JE, et al. The ORBIT bleeding score: a simple bedside score to assess bleeding risk in atrial fibrillation. Eur Heart J. 2015;36(46):3258-64.
64. Hart RG, Boop BS, Anderson DC. Oral Anticoagulants and Intracranial Hemorrhage. Stroke. 1995;26(8):1471-7.
65. Paulus E, Komperda K, Park G, Fusco J. Anticoagulation Therapy Considerations in Factor VII Deficiency. Drug Safety - Case Reports. 2016;3.
66. Ageno W, Gallus AS, Wittkowsky A, Crowther M, Hylek EM, Palareti G. Oral anticoagulant therapy: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest. 2012;141(2 Suppl):e44S-e88S.
67. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med. 2007;146(12):857-67.
68. Hirsh J, Fuster V, Ansell J, Halperin JL. American Heart Association/American College of Cardiology Foundation guide to warfarin therapy. J Am Coll Cardiol. 2003;41(9):1633-52.
69. Hasan SS, Kow CS, Curley LE, Baines DL, Babar ZU. Economic evaluation of prescribing conventional and newer oral anticoagulants in older adults. Expert Rev Pharmacoecon Outcomes Res. 2018;18(4):371-7.

70. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, et al. Dabigatran versus Warfarin in Patients with Atrial Fibrillation. New England Journal of Medicine. 2009;361(12):1139-51.
71. 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.
72. Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL, et al. Edoxaban versus Warfarin in Patients with Atrial Fibrillation. New England Journal of Medicine. 2013;369(22):2093-104.
73. Administration UFD. Pradaxa (dabigatran) Medical Review, NDA 22-512. 2010.
74. T P. Letter 80, Dabigatran for atrial fibrillation Why we can not rely on RE-LY. Therapeutics Initiative. 2011(March 2011).
75. Christopoulos C. Questions about the reliability of ROCKET-AF anticoagulation data. BMJ. 2016;352:i1039.
76. Cohen D. Rivaroxaban: can we trust the evidence? BMJ. 2016;352:i575.
77. Kmietowicz Z. Boehringer Ingelheim withheld safety analyses on new anticoagulant, The BMJ investigation finds. BMJ: British Medical Journal. 2014;349.
78. Mavrakanas T, Bounameaux H. The potential role of new oral anticoagulants in the prevention and treatment of thromboembolism. Pharmacol Ther. 2011;130(1):46-58.
79. Shah M, Avgil Tsadok M, Jackevicius CA, Essebag V, Eisenberg MJ, Rahme E, et al. Warfarin use and the risk for stroke and bleeding in patients with atrial fibrillation undergoing dialysis. Circulation. 2014;129(11):1196-203.
80. Afzal S, Zaidi STR, Merchant HA, Babar ZU, Hasan SS. Prescribing trends of oral anticoagulants in England over the last decade: a focus on new and old drugs and adverse events reporting. J Thromb Thrombolysis. 2021;52(2):646-53.
81. Agency EM. Assessment report Eliquis apixaban. Procedure No.:EMEA/H/C/002148/X/04/G. EMA/641505/2012. In: (CHMP) CfMPfHU, editor. 2012.
82. Yao X, Abraham NS, Sangaralingham LR, Bellolio MF, McBane RD, Shah ND, et al. Effectiveness and Safety of Dabigatran, Rivaroxaban, and Apixaban Versus Warfarin in Nonvalvular Atrial Fibrillation. J Am Heart Assoc. 2016;5(6).
83. Lip GY, Keshishian A, Kamble S, Pan X, Mardekian J, Horblyuk R, et al. Real-world comparison of major bleeding risk among non-valvular atrial fibrillation patients initiated on apixaban, dabigatran, rivaroxaban, or warfarin. A propensity score matched analysis. Thromb Haemost. 2016;116(5):975-86.
84. Lip GY, Pan X, Kamble S, Kawabata H, Mardekian J, Masseria C, et al. Major bleeding risk among non-valvular atrial fibrillation patients initiated on apixaban, dabigatran, rivaroxaban or warfarin: a "real-world" observational study in the United States. Int J Clin Pract. 2016;70(9):752-63.
85. Coleman CI, Antz M, Bowrin K, Evers T, Simard EP, Bonnemeier H, et al. Real-world evidence of stroke prevention in patients with nonvalvular atrial fibrillation in the United States: the REVISIT-US study. Curr Med Res Opin. 2016;32(12):2047-53.
86. Li XS, Deitelzweig S, Keshishian A, Hamilton M, Horblyuk R, Gupta K, et al. Effectiveness and safety of apixaban versus warfarin in non-valvular atrial fibrillation patients in "real-world" clinical practice. A propensity-matched analysis of 76,940 patients. Thromb Haemost. 2017;117(6):1072-82.
87. Coleman CI, Peacock WF, Bunz TJ, Alberts MJ. Effectiveness and Safety of Apixaban, Dabigatran, and Rivaroxaban Versus Warfarin in Patients With Nonvalvular Atrial Fibrillation and Previous Stroke or Transient Ischemic Attack. Stroke. 2017;48(8):2142-9.
88. Lin J, Trocio J, Gupta K, Mardekian J, Lingohr-Smith M, Menges B, et al. Major bleeding risk and healthcare economic outcomes of non-valvular atrial fibrillation patients

newly-initiated with oral anticoagulant therapy in the real-world setting. J Med Econ. 2017;20(9):952-61.
89. Adeboyeje G, Sylwestrzak G, Barron JJ, White J, Rosenberg A, Abarca J, et al. Major Bleeding Risk During Anticoagulation with Warfarin, Dabigatran, Apixaban, or Rivaroxaban in Patients with Nonvalvular Atrial Fibrillation. Journal of Managed Care & Specialty Pharmacy. 2017;23(9):968-78.
90. Deitelzweig S, Luo X, Gupta K, Trocio J, Mardekian J, Curtice T, et al. Comparison of effectiveness and safety of treatment with apixaban vs. other oral anticoagulants among elderly nonvalvular atrial fibrillation patients. Curr Med Res Opin. 2017;33(10):1745-54.
91. Hernandez I, Zhang Y, Saba S. Comparison of the Effectiveness and Safety of Apixaban, Dabigatran, Rivaroxaban, and Warfarin in Newly Diagnosed Atrial Fibrillation. Am J Cardiol. 2017;120(10):1813-9.
92. Lopes RD, Fusco MD, Keshishian A, Luo X, Li X, Masseria C, et al. Abstract 16487: Effectiveness and Safety of Apixaban versus Other Oral Anticoagulants in Older Adults With Non-Valvular Atrial Fibrillation and Concomitant Coronary Artery Disease or Peripheral Arterial Disease. Circulation. 2017;136(suppl_1):A16487-A.
93. Li X, Keshishian A, Hamilton M, Horblyuk R, Gupta K, Luo X, et al. Apixaban 5 and 2.5 mg twice-daily versus warfarin for stroke prevention in nonvalvular atrial fibrillation patients: Comparative effectiveness and safety evaluated using a propensity-score-matched approach. PLoS One. 2018;13(1):e0191722.
94. Amin A, Keshishian A, Vo L, Zhang Q, Dina O, Patel C, et al. Real-world comparison of all-cause hospitalizations, hospitalizations due to stroke and major bleeding, and costs for non-valvular atrial fibrillation patients prescribed oral anticoagulants in a US health plan. J Med Econ. 2018;21(3):244-53.
95. Lopes RD, Steffel J, Di Fusco M, Keshishian A, Luo X, Li X, et al. Effectiveness and Safety of Anticoagulants in Adults with Non-valvular Atrial Fibrillation and Concomitant Coronary/Peripheral Artery Disease. Am J Med. 2018;131(9):1075-85.e4.
96. Amin A, Keshishian A, Trocio J, Dina O, Le H, Rosenblatt L, et al. A Real-World Observational Study of Hospitalization and Health Care Costs Among Nonvalvular Atrial Fibrillation Patients Prescribed Oral Anticoagulants in the U.S. Medicare Population. J Manag Care Spec Pharm. 2018;24(9):911-20.
97. Lip GYH, Keshishian A, Li X, Hamilton M, Masseria C, Gupta K, et al. Effectiveness and Safety of Oral Anticoagulants Among Nonvalvular Atrial Fibrillation Patients. Stroke. 2018;49(12):2933-44.
98. Gupta K, Trocio J, Keshishian A, Zhang Q, Dina O, Mardekian J, et al. Effectiveness and safety of direct oral anticoagulants compared to warfarin in treatment naïve non-valvular atrial fibrillation patients in the US Department of defense population. BMC Cardiovasc Disord. 2019;19(1):142.
99. Amin A, Keshishian A, Trocio J, Dina O, Le H, Rosenblatt L, et al. A Real-World Observational Study of Hospitalization and Health Care Costs Among Nonvalvular Atrial Fibrillation Patients Prescribed Oral Anticoagulants in the U.S. Medicare Population. J Manag Care Spec Pharm. 2020;26(5):639-51.
100. Franklin J BaWsH. Replication of the ARISTOTLE Anticoagulant Trial in Healthcare Claims Data. Available from: https://www.clinicaltrials.gov/ct2/show/NCT045930302020 [
101. Larsen TB, Skjoth F, Nielsen PB, Kjaeldgaard JN, Lip GYH, T.B. L, et al. Comparative effectiveness and safety of non-Vitamin K antagonist oral anticoagulants and warfarin in patients with Atrial fibrillation: Propensity weighted nationwide cohort study. BMJ (Online). 2016;353:lart032161.
102. Lamberts M, Staerk L, Olesen JB, Fosbøl EL, Hansen ML, Harboe L, et al. Major Bleeding Complications and Persistence With Oral Anticoagulation in Non-Valvular Atrial

Fibrillation: Contemporary Findings in Real-Life Danish Patients. Journal of the American Heart Association. 2017;6(2):e004517.
103. Nielsen PB, Skjøth F, Søgaard M, Kjældgaard 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.
104. Staerk L, Fosbøl EL, Lip GYH, Lamberts M, Bonde AN, Torp-Pedersen C, et al. Ischaemic and haemorrhagic stroke associated with non-vitamin K antagonist oral anticoagulants and warfarin use in patients with atrial fibrillation: a nationwide cohort study. Eur Heart J. 2017;38(12):907-15.
105. Coleman CI, Antz M. Real-world evidence with apixaban for stroke prevention in patients with nonvalvular atrial fibrillation in Germany: a retrospective study (REASSESS). Intern Emerg Med. 2017;12(3):419-22.
106. Hohnloser SH, Basic E, Nabauer M. Comparative risk of major bleeding with new oral anticoagulants (NOACs) and phenprocoumon in patients with atrial fibrillation: a post-marketing surveillance study. Clin Res Cardiol. 2017;106(8):618-28.
107. Lip GYH, Skjøth F, Nielsen PB, Kjældgaard JN, Larsen TB. Effectiveness and Safety of Standard-Dose Nonvitamin K Antagonist Oral Anticoagulants and Warfarin Among Patients With Atrial Fibrillation With a Single Stroke Risk Factor: A Nationwide Cohort Study. JAMA Cardiol. 2017;2(8):872-81.
108. Coleman CI, Peacock WF, Antz M. Comparative Effectiveness and Safety of Apixaban and Vitamin K Antagonist Therapy in Patients With Nonvalvular Atrial Fibrillation Treated in Routine German Practice. Heart Lung Circ. 2018;27(3):390-3.
109. Hohnloser SH, Basic E, Hohmann C, Nabauer M. Effectiveness and Safety of Non-Vitamin K Oral Anticoagulants in Comparison to Phenprocoumon: Data from 61,000 Patients with Atrial Fibrillation. Thromb Haemost. 2018;118(3):526-38.
110. Själander S, Sjögren V, Renlund H, Norrving B, Själander A. Dabigatran, rivaroxaban and apixaban vs. high TTR warfarin in atrial fibrillation. Thromb Res. 2018;167:113-8.
111. Vinogradova Y, Coupland C, Hill T, Hippisley-Cox J. Risks and benefits of direct oral anticoagulants versus warfarin in a real world setting: cohort study in primary care. Bmj. 2018;362:k2505.
112. Ramagopalan S, Allan V, Saragoni S, Esposti LD, Alessandrini D, Perrone V, et al. Patient characteristics and bleeding events in nonvalvular atrial fibrillation patients treated with apixaban or vitamin K antagonists: real-world evidence from Italian administrative databases. J Comp Eff Res. 2018;7(11):1063-71.
113. Marietta M, Banchelli F, Pavesi P, Manotti C, Quintavalla R, Sinigaglia T, et al. Direct oral anticoagulants vs non-vitamin K antagonist in atrial fibrillation: A prospective, propensity score adjusted cohort study. European Journal of Internal Medicine. 2019;62:9-16.
114. Ramagopalan SV, Sicras-Mainar A, Polanco-Sanchez C, Carroll R, de Bobadilla JF. Patient characteristics and stroke and bleeding events in nonvalvular atrial fibrillation patients treated with apixaban and vitamin K antagonists: a Spanish real-world study. J Comp Eff Res. 2019;8(14):1201-12.
115. Van Ganse E, Danchin N, Mahé I, Hanon O, Jacoud F, Nolin M, et al. Comparative Safety and Effectiveness of Oral Anticoagulants in Nonvalvular Atrial Fibrillation: The NAXOS Study. Stroke. 2020;51(7):2066-75.
116. Warkentin L, Klohn F, Deiters B, Kühlein T, Hueber S. Vitamin-K-antagonist phenprocoumon versus direct oral anticoagulants in patients with atrial fibrillation: a real-world analysis of German claims data. BMJ Open. 2023;13(1):e063490.

117. Le Heuzey JY, Ammentorp B, Darius H, De Caterina R, Schilling RJ, Schmitt J, et al. Differences among western European countries in anticoagulation management of atrial fibrillation. Data from the PREFER IN AF registry. Thromb Haemost. 2014;111(5):833-41.
118. Kohsaka S, Murata T, Izumi N, Katada J, Wang F, Terayama Y. Bleeding risk of apixaban, dabigatran, and low-dose rivaroxaban compared with warfarin in Japanese patients with non-valvular atrial fibrillation: a propensity matched analysis of administrative claims data. Curr Med Res Opin. 2017;33(11):1955-63.
119. Cha MJ, Choi EK, Han KD, Lee SR, Lim WH, Oh S, et al. Effectiveness and Safety of Non-Vitamin K Antagonist Oral Anticoagulants in Asian Patients With Atrial Fibrillation. Stroke. 2017;48(11):3040-8.
120. Chan YH, See LC, Tu HT, Yeh YH, Chang SH, Wu LS, et al. Efficacy and Safety of Apixaban, Dabigatran, Rivaroxaban, and Warfarin in Asians With Nonvalvular Atrial Fibrillation. J Am Heart Assoc. 2018;7(8).
121. Kohsaka S, Katada J, Saito K, Terayama Y. Safety and effectiveness of apixaban in comparison to warfarin in patients with nonvalvular atrial fibrillation: a propensity-matched analysis from Japanese administrative claims data. Curr Med Res Opin. 2018;34(9):1627-34.
122. Kohsaka S, Katada J, Saito K, Jenkins A, Li B, Mardekian J, et al. Safety and effectiveness of non-vitamin K oral anticoagulants versus warfarin in real-world patients with non-valvular atrial fibrillation: a retrospective analysis of contemporary Japanese administrative claims data. Open Heart. 2020;7(1):e001232.
123. Steinberg BA, Shrader P, Thomas L, Ansell J, Fonarow GC, Gersh BJ, et al. Off-Label Dosing of Non-Vitamin K Antagonist Oral Anticoagulants and Adverse Outcomes: The ORBIT-AF II Registry. J Am Coll Cardiol. 2016;68(24):2597-604.
124. Harrington J, Arps K, Wu A, Carnicelli AP, Chiswell K, Chrischilles E, et al. Reduced dose, but not reduced risk: rates of inappropriate apixaban dose reduction and stroke and bleeding incidence. European Heart Journal. 2022;43(Supplement_2).
125. Brown JD, Shewale AR, Talbert JC. Adherence to Rivaroxaban, Dabigatran, and Apixaban for Stroke Prevention for Newly Diagnosed and Treatment-Naive Atrial Fibrillation Patients: An Update Using 2013-2014 Data. J Manag Care Spec Pharm. 2017;23(9):958-67.
126. Levintow SN, Nielson CM, Hernandez RK, Breskin A, Pritchard D, Lash TL, et al. Pragmatic considerations for negative control outcome studies to guide non-randomized comparative analyses: A narrative review. Pharmacoepidemiology and Drug Safety. 2023;32(6):599-606.
127. Shen AY, Yao JF, Brar SS, Jorgensen MB, Chen W. Racial/ethnic differences in the risk of intracranial hemorrhage among patients with atrial fibrillation. J Am Coll Cardiol. 2007;50(4):309-15.
128. Wang KL, Lip GY, Lin SJ, Chiang CE. Non-Vitamin K Antagonist Oral Anticoagulants for Stroke Prevention in Asian Patients With Nonvalvular Atrial Fibrillation: Meta-Analysis. Stroke. 2015;46(9):2555-61.
129. Cotté FE, Benhaddi H, Duprat-Lomon I, Doble A, Marchant N, Letierce A, et al. Vitamin K antagonist treatment in patients with atrial fibrillation and time in therapeutic range in four European countries. Clin Ther. 2014;36(9):1160-8.
130. Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L. Real-world effects of medications for chronic obstructive pulmonary disease: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results. 2018;8:e019475.
131. Rose M BN. Center for Drug Evaluation and Research Application number: 202155Orig1s000 MEDICAL REVIEW(S).

132. EMA. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) Assessment report Eliquis apixaban. Procedure No.:EMEA/H/C/002148/X/04/G. EMA/641505/2012, 2012, page 35.
133. Vinereanu D, Stevens SR, Alexander JH, Al-Khatib SM, Avezum A, Bahit MC, et al. Clinical outcomes in patients with atrial fibrillation according to sex during anticoagulation with apixaban or warfarin: a secondary analysis of a randomized controlled trial. Eur Heart J. 2015;36(46):3268-75.
134. Halvorsen S, Atar D, Yang H, De Caterina R, Erol C, Garcia D, et al. Efficacy and safety of apixaban compared with warfarin according to age for stroke prevention in atrial fibrillation: observations from the ARISTOTLE trial. Eur Heart J. 2014;35(28):1864-72.
135. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol. 2003;158(9):915-20.
136. Webster-Clark M, Mavros P, Garry EM, Stürmer T, Shmuel S, Young J, et al. Alternative analytic and matching approaches for the prevalent new-user design: A simulation study. Pharmacoepidemiology and Drug Safety. 2022;31(7):796-803.
137. Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. Pharmacoepidemiol Drug Saf. 2017;26(4):459-68.
138. Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol. 2008;167(4):492-9.
139. Wallentin L, Lopes RD, Hanna M, Thomas L, Hellkamp A, Nepal S, et al. Efficacy and Safety of Apixaban Compared With Warfarin at Different Levels of Predicted International Normalized Ratio Control for Stroke Prevention in Atrial Fibrillation. Circulation. 2013;127(22):2166-76.
140. Wieloch M, Själander A, Frykman V, Rosenqvist M, Eriksson N, Svensson PJ. Anticoagulation control in Sweden: reports of time in therapeutic range, major bleeding, and thrombo-embolic complications from the national quality registry AuriculA. European Heart Journal. 2011;32(18):2282-9.
141. Dlott JS, George RA, Huang X, Odeh M, Kaufman HW, Ansell J, et al. National Assessment of Warfarin Anticoagulation Therapy for Stroke Prevention in Atrial Fibrillation. Circulation. 2014;129(13):1407-14.
142. Goto S, Zhu J, Liu L, Oh BH, Wojdyla DM, Aylward P, et al. Efficacy and safety of apixaban compared with warfarin for stroke prevention in patients with atrial fibrillation from East Asia: a subanalysis of the Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (ARISTOTLE) Trial. Am Heart J. 2014;168(3):303-9.
143. Sabir I, Khavandi K, Brownrigg J, Camm AJ. Oral anticoagulants for Asian patients with atrial fibrillation. Nat Rev Cardiol. 2014;11(5):290-303.
144. Gillott RG, Willan K, Kain K, Sivananthan UM, Tayebjee MH. South Asian ethnicity is associated with a lower prevalence of atrial fibrillation despite greater prevalence of established risk factors: a population-based study in Bradford Metropolitan District. Europace. 2017;19(3):356-63.
145. Dell'Aniello S, Renoux C, Suissa S. Regarding: Webster-Clark M., et al. Alternative analytic and matching approaches for the prevalent new-user design: a simulation study. Pharmacoepidemiol Drug Saf, 31, pp. 796–803; 2022. Pharmacoepidemiology and Drug Safety. 2022;31(12):1317-8.
146. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology. 2009;20(4):512-22.

147. Wyss R, Fireman B, Rassen JA, Schneeweiss S. Erratum: High-dimensional Propensity Score Adjustment in Studies of Treatment Effects Using Health Care Claims Data. Epidemiology. 2018;29(6):e63-e4.
148. Tazare J, Wyss R, Franklin JM, Smeeth L, Evans SJW, Wang SV, et al. Transparency of high-dimensional propensity score analyses: Guidance for diagnostics and reporting. Pharmacoepidemiol Drug Saf. 2022;31(4):411-23.
149. Zubizarreta JR PR, Rosenbaum PR. Matching for balance, pairing for heterogeneity in an observational study of the effectiveness of for-profit and not-for-profit high schools in Chile. The Annals of Applied Statistics. 2014;8(1).
150. Visconti G ZJ. Handling limited overlap in observational studies with cardinality matching. . Observational Studies 2018;4: 217-249.
151. Fortin SP, Johnston SS, Schuemie MJ. Applied comparison of large-scale propensity score matching and cardinality matching for causal inference in observational research. BMC Med Res Methodol. 2021;21(1):109.
152. Cheung CL, Sing CW, Lau WCY, Li GHY, Lip GYH, Tan KCB, et al. Treatment with direct oral anticoagulants or warfarin and the risk for incident diabetes among patients with atrial fibrillation: a population-based cohort study. Cardiovasc Diabetol. 2021;20(1):71.
153. Huang HK, Liu PP, Lin SM, Hsu JY, Peng CC, Munir KM, et al. Risk of developing diabetes in patients with atrial fibrillation taking non-vitamin K antagonist oral anticoagulants or warfarin: A nationwide cohort study. Diabetes Obes Metab. 2021;23(2):499-507.
154. Ho HJ, Komai M, Shirakawa H. Beneficial Effects of Vitamin K Status on Glycemic Regulation and Diabetes Mellitus: A Mini-Review. Nutrients. 2020;12(8).
155. Friberg L, Rosenqvist M, Lip GYH. Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. European Heart Journal. 2012;33(12):1500-10.
156. Lip GY, Frison L, Halperin JL, Lane DA. Identifying patients at high risk for stroke despite anticoagulation: a comparison of contemporary stroke risk stratification schemes in an anticoagulated atrial fibrillation cohort. Stroke. 2010;41(12):2731-8.
157. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. 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;137(2):263-72.
158. Abadie A, Spiess J. Robust Post-Matching Inference. Journal of the American Statistical Association. 2022;117(538):983-95.
159. Austin PC, Small DS. The use of bootstrapping when using propensity-score matching without replacement: a simulation study. Stat Med. 2014;33(24):4306-19.

Appendices

Appendix 1

A1.1 Supplementary material from Chapter 2 Background

Table A1.1.1 Annual risk of thromboembolism by CHA_2DS_2-VASc score

	Annual risk of thromboembolism (%), adjusted to remove effect of treatment	
CHA_2DS_2-VASc score	**Friberg 2012(155)[a]**	**NICE[b]**
0	0.3	0
1	1.0	1.3
2	3.3	2.2
3	5.3	3.2
4	7.8	4.0
5	11.7	6.7
6	15.9	9.8
7	18.4	9.6
8	17.9	6.7
9	20.3	15.2

Thromboembolism includes stroke, TIA, or systemic embolism with adjusted risk an estimate of the risk in untreated patients (ie receiving no anticoagulation or aspirin treatment).

a Friberg 2012 results from the Swedish Atrial Fibrillation cohort study with raw rates adjusted assuming that aspirin provided a 22% reduction in risk.

b NICE risks reproduced form the clinical knowledge summary CHA_2DS_2-VASc clinical risk estimation for stroke or other thromboembolic events and based on the results of Lip et al(156), with adjustment assuming that warfarin provides a 64% reduction in risk of thromboembolism Hart et al (67).

Table A1.1.2 ORBIT bleeding risk scoring system

Risk factor		Points
	Male with haemoglobin <130 g/L or haematocrit <40%	2
	Female with haemoglobin <120 g/L or haematocrit <36%	2
	History of bleeding (eg, gastrointestinal or intracranial bleeding, or haemorrhagic stroke)	2
	Age > 74 years	1
	Estimated glomerular filtration rate < 60 mL/min/1.73m^2	1
	Treated with antiplatelets	1

Table A1.1.3 Bleeding risk by ORBIT score from O'Brien et al (63)

ORBIT bleeding score	Major bleeds per 100 patient-years (95% CI)	ORBIT bleeding score category: major bleeds per 100 patient-years
0	1.7 (1.2, 2.4)	Low (0-2): 2.4
1	2.3 (1.9, 2.9)	
2	2.9 (2.3, 3.5)	
3	4.7 (4.0, 5.6)	Medium (3): 4.7
4	6.8 (5.8, 8.1)	High (≥4): 8.1
5	9.0 (7.2, 11.2)	
6	12.3 (9.0, 16.7)	
7	14.9 (8.9, 25.3)	

Observed bleeding rates from ORBIT-AF (Outcomes registry for better informed treatment).
Table from: O'Brien EC et al. The ORBIT bleeding score: a simple bedside score to assess bleeding risk in atrial fibrillation. Eur Heart J. 2015 Dec 7;36(46):3258-64. doi: 10.1093/eurheartj/ehv476.
Table 3 Outcomes registry for better informed treatment bleeding risk score and observed major bleeding rates (63)

Appendix 2

A2.1 Supplementary material from Research Paper 1: BMJ Open protocol

Appendix Table: ARISTOTLE Inclusion and Exclusion Criteria Algorithms for EHR

To be trial eligible a patient must have all inclusion criteria (IE01 to IE03)=Y and no exclusion (IE05 to IE27c)=Y

Criteria #	Used?	Criteria Text (from ARISTOTLE protocol)	Implementation Rule and Notes
		Inclusion Critera (IE01 to IE04a)	
IE01	Y	Age ≥ 18 years	Calculate age at index date, day and month of birth not available therefore calculate age by assuming birthdate=01-July-birthyear: age =(indexdate-birthdate)/365.25 If age ge 18 then IE01=Y.
IE02	Y	In atrial fibrillation or atrial flutter not due to a reversible cause and documented by ECG at the time of enrollment. OR If not in atrial fibrillation/flutter at the time of enrollment, must have atrial fibrillation/flutter documented on two separate occasions, not due to a reversible cause at least 2 weeks apart in the 12 months prior to enrollment. Atrial fibrillation/flutter may be documented by ECG, or as an episode lasting at least one minute on a rhythm strip, Holter recording, or intracardiac electrogram (from an implanted pacemaker or defibrillator).	If patient has medical record corresponding to atrial fibrillation or atrial flutter on or prior to index date then IE02=Y.
		One or more of the following risk factor(s) for stroke:	IE03=Y if at least one of (IE03a, IE03b, IE03c, IE03d, IE03e) is Y.
IE03a	Y	Age 75 years or older	See IE01 for derivation of age at index date. If age ge 75 then IE03a=Y
IE03b	Y	Prior stroke, TIA or systemic embolus	If patient has medical record corresponding to stroke, TIA, or systemic embolus diagnosis on or prior to index date then IE03b=Y. Codelist search terms include 'stroke', 'cerebrovascular accident', 'cerebral infarction', 'lacunar', 'transient ischaemic attack', and synonyms for these.
IE03c	Y	Either symptomatic congestive heart failure within 3 months or left ventricular dysfunction with an LV ejection fraction (LVEF) ≤ 40% by echocardiography, radionuclide study or contrast angiography	If patient has medical record corresponding to congestive heart failure or left ventricular dysfunction diagnosis on or prior to index date then IE03c=Y. Codelist search terms include 'heart failure', 'cardiac failure', 'congestive heart failure', 'cardiomyopathy', 'left ventricular dysfunction', 'left ventricular', 'lvef', 'new york heart association classification', 'hypertensive heart', and synonyms for these.
IE03d	Y	Diabetes mellitus	If patient has medical record corresponding to diabetes diagnosis on or prior to index date then IE03d=Y. Codelist search terms include 'diabetes', both type 1 and type 2 diabetes are included.

Criteria #	Used?	Criteria Text (from ARISTOTLE protocol)	Implementation Rule and Notes
IE03e	Y	Hypertension requiring pharmacological treatment	If patient has medical record corresponding to hypertension on or prior to index date AND a prescription for an antihypertensive on or prior to index date then IE03e=Y. Hypertension codelist search terms include 'hyperten', 'high blood pressure', 'nephrosclerosis', and synonyms for these.
IE04	N	Women of childbearing potential (WOCBP) must be using an adequate method of contraception to avoid pregnancy throughout the treatment period of the study or for 2 weeks after the last dose of study medication, whichever is longer, in such a manner that the risk of pregnancy is minimized. WOCBP must have a negative serum or urine pregnancy test (minimum sensitivity 25 IU/L or equivalent units of HCG) within 48 hours prior to the start of investigational product.	This criteria is only partially applied - women with evidence of recent pregnancy or breastfeeding will be excluded (see IE27c).
IE04b	N	All subjects must provide signed written informed consent.	N/A for observational study
		Exclusion criteria (IE05 to IE27d)	
IE05	Y	Atrial fibrillation or flutter due to reversible causes (e.g. thyrotoxicosis, pericarditis)	If patient has medical record corresponding to reversible AF causes on or prior to index date then IE05=Y. Codelist search terms include 'thyrotoxicosis', 'pericarditis', and synonyms for these.
IE06	Y	Clinically significant (moderate or severe) mitral stenosis	If patient has medical record corresponding to mitral stenosis on or prior to index date then IE06=Y. Cannot determine clinical significance of 'mitral stenosis' terms in CPRD therefore assume if there is a record of mitral stenosis the condition is clinically significant.
IE07	Y	Increased bleeding risk that is believed to be a contraindication to oral anticoagulation (e.g. previous intracranial hemorrhage)	If patient has medical record corresponding to increased bleeding risk on or prior to index date then IE07=Y. Codelist search terms include 'haemorrhag', 'bleed', 'aneurysm', (('intracranial' or 'brain') and ('neoplasm' or 'tumour' or 'cancer')), 'arteriovenous malformation', 'immune thrombocytopenic purpura', 'evans disease', 'hemolytic anemia', 'haemophilia', 'von willebrand disease', ('glanzmann' and 'thrombasthenia'), 'wiskott–aldrich syndrome', 'thrombocytopenia' and synonyms for these. For some forms of more common past bleeding event such as bleeding related to menstrual or uterine bleeding, bleeding associated with surgery or injury, bleeding associated with ulcer or gastritis, eye bleeding (retinal, conjunctival) we apply the additional criteria that these must be within the last two years to be included as evidence of increased bleeding risk.

Powell EM, *et al. BMJ Open* 2021; 11:e042947. doi: 10.1136/bmjopen-2020-042947

Criteria #	Used?	Criteria Text (from ARISTOTLE protocol)	Implementation Rule and Notes
IE08	Y	Conditions other than atrial fibrillation that require chronic anticoagulation (e.g. prosthetic mechanical heart valve)	If patient has medical record corresponding to a condition other than atrial fibrillation that requires chronic anticoagulation on or prior to index date then IE08=Y. Codelist search terms include (('heart' or 'valve') and ('prosth' or 'mechanical')), 'venous thromb', and synonyms for these.
IE09	Y	Persistent, uncontrolled hypertension (systolic BP > 180 mm Hg, or diastolic BP > 100 mm Hg)	If patient has at least 2 blood pressure readings over the limit (systolic BP > 180 mm Hg, or diastolic BP > 100 mm Hg) in the 6 months prior to the index date OR the patient has a medical record (within 180 days prior to index date) indicating uncontrolled hypertension then IE09=Y Codelist search terms include 'poor hypertension control', 'hypertensive crisis', 'malignant hypertension', 'severe hypertension', 'hypertension resistant to drug therapy', and synonyms for these.
IE10	Y	Active infective endocarditis	If patient has medical record corresponding to endocarditis on or prior to index date then IE10=Y.
IE11	N	Planned major surgery	N/A – do not look at future events when determining eligibility
IE12	N	Planned atrial fibrillation or flutter ablation procedure	N/A – do not look at future events when determining eligibility
IE13	N	Use of an unapproved, investigational drug or device within the past 30 days	N/A – not appropriate to apply when looking at observational data
IE14	Y	Required treatment with aspirin > 165 mg/day	If patient has a prescription for aspirin with dose > 165 mg/day and prescription data suggests drug exposure ongoing at index date then IE14=Y. Note this will not pick up patients taking regular aspirin over the counter (study limitation).
IE15	Y	Simultaneous treatment with both aspirin and a thienopyridine (e.g., clopidogrel, ticlopidine)	If both aspirin and thienopyridine ongoing at index date (ie derived exposure covers index date) then IE15=Y.
IE16	Y	Severe comorbid condition with life expectancy of ≤ 1 year	If patient has medical record corresponding to a condition with a low median survival time then IE16=Y. Codelist search terms include pancreatic, oesophageal, stomach, liver, gallbladder, biliary duct, bladder, lung or brain cancer, multiple myeloma, mesothelioma, CJD, and synonyms for these.
IE17	Y	Active alcohol or drug abuse, or psychosocial reasons that make study participation impractical	If patient has medical record corresponding to drug or alcohol abuse or any complications of abuse, conditions involving an impaired mental state (dementia including subtypes such as Alzheimer's), severe mental health conditions (schizophrenia, psychosis, bipolar) then IE17=Y.
IE18	Y	Recent ischemic stroke (within 7 days)	If patient has medical record corresponding to ischemic stroke within 7 days of index date (prior) then IE18=Y.

Criteria #	Used?	Criteria Text (from ARISTOTLE protocol)	Implementation Rule and Notes
IE19	Y	Severe renal insufficiency (serum creatinine > 2.5 mg/dL or a calculated creatinine clearance < 25 mL/min, See Section 6.3.2.2)	If patient has lab result showing serum creatinine > 2.5 mg/dL or a calculated creatinine clearance < 25 mL/min within 90 days prior to index date OR a medical record corresponding to severe renal insufficiency (chronic kidney disease stage 4 or 5, dialysis) then IE19=Y
IE20	Y	ALT or AST > 2X ULN or a Total Bilirubin ≥ 1.5X ULN (unless an alternative causative factor [e.g., Gilbert's syndrome] is identified)	If patient has lab result showing ALT or AST > 2X ULN or a Total Bilirubin ≥ 1.5X ULN within 90 days prior to index date (AND no diagnosis of Gilbert's syndrome) then IE20=Y
IE21	Y	Platelet count ≤ 100,000/ mm3	If patient has lab result showing platelet count ≤ 100,000/ mm3 within 90 days prior to index date OR a medical record of thrombocytopenia within 90 days prior to index date then IE21=Y
IE22	Y	Hemoglobin < 9 g/dL	If patient has lab result showing hemoglobin < 9 g/dL within 90 days prior to index date then IE22=Y
IE23	N	Inability to comply with INR monitoring	Patients unlikely to be able to comply with INR monitoring – evidence of drug or alcohol abuse, impaired mental state, severe mental health conditions. All these conditions are already excluded by IE17
IE24	N	Prior randomization into an apixaban clinical study	N/A
IE25	N	Prisoners or subjects who are involuntarily incarcerated	N/A
IE26	N	Subjects who are compulsorily detained for treatment of either a psychiatric or physical (e.g., infectious disease) illness	N/A
	N	Women of child bearing potential (WOCBP) unwilling or unable to use an acceptable method to avoid pregnancy:	N/A – see IE27c
IE27a	N	WOCBP using a prohibited contraceptive method	N/A

Criteria #	Used?	Criteria Text (from ARISTOTLE protocol)	Implementation Rule and Notes
IE27b	N	WOCBP include any female who has experienced menarche and who has not undergone successful surgical sterilization (hysterectomy, bilateral tubal ligation, or bilateral oophorectomy) or is not postmenopausal [defined as amenorrhea ≥ 12 consecutive months, or women on hormone replacement therapy (HRT) with documented serum follicle stimulating hormone (FSH) level > 35 mIU/mL]. Even women who are using oral contraceptives, other hormonal contraceptives (vaginal products, skin patches, or implanted or injectable products), or mechanical products such as an intrauterine device or barrier methods (diaphragm, condoms, spermicides) to prevent pregnancy, or are practicing abstinence or where their partner is sterile (e.g., vasectomy) should be considered to be of child bearing potential	N/A
IE27c	Y	Women who are pregnant or breastfeeding	Exclude women who have any medical codes relating to pregnancy (regardless of the outcome of the pregnancy), childbirth, antenatal or postnatal care, or breastfeeding in the 3 years prior to the patient's index date.
IE27d	N	Women with a positive pregnancy test on enrollment or prior to administration of investigational product.	N/A – covered by IE27c

Note: Algorithms are under development as part of this study and may be further refined prior to being finalised.
N/A = Not Applicable. For IE19-IE22 involving lab results a pragmatic approach will be taken in which a patient is assumed not to have the exclusion criteria if there is no lab result available in the 90 days prior to index date and the latest available lab result prior to index date does not meet the criteria.

Powell EM, *et al. BMJ Open* 2021; 11:e042947. doi: 10.1136/bmjopen-2020-042947

A2.2 Codelists used in the project

Full codelists available to view and download from

https://datacompass.lshtm.ac.uk/id/eprint/3590/

Snomed codes for Atrial Fibrillation for CPRD Aurum

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
260825100000011	6 15964901000119107	3322864015	atypical atrial flutter
986894100000611	6 15964901000119107	3323015018	atrial flutter type 2
986895100000611	9 15964901000119107	3323016017	atrial flutter type ii
256478018	164889003	256478018	ecg: atrial fibrillation
458661100000611	2 164889003	3300048013	electrocardiographic atrial fibrillation
458663100000611	8 164890007	3300081014	electrocardiographic atrial flutter
458664100000611	1 164890007	3300082019	ekg: atrial flutter
256479014	164890007	256479014	ecg: atrial flutter
783761000006118	175146007	271235013	implant intravenous pacemaker for atrial fibrillation
865521000006116	175146007	865521000006116	pacer controlled atrial fibril
182395100000611	1 1823951000006107	1823951000006111	atrial fibrillation confirmed
182405100000611	3 1824051000006109	1824051000006113	atrial fibrillation clinical pathway protocol followed
185643100000611	8 1856431000006102	1856431000006118	atrial fibrillation follow-up
193202100000611	7 1932021000006101	1932021000006117	3d study - problems with atrial fibrillation management
300130013	195080001	300130013	atrial fibrillation and flutter
300132017	195080001	300130013	atrial fibrillation and flutter nos
350465014	233911009	350465014	non-rheumatic atrial fibrillation
505728100000611	7 233911009	350466010	nraf - non-rheumatic atrial fibrillation
406861000000119	248411000000105	406861000000119	atrial fibrillation annual review
566961100000611	9 282825002	421234013	intermittent atrial fibrillation
566959100000611	3 282825002	421232012	af - paroxysmal atrial fibrillation
566960100000611	7 282825002	421233019	paf - paroxysmal atrial fibrillation
421235014	282825002	421235014	paroxysmal atrial fibrillation
601640100000611	5 312442005	2986292016	history of atrial fibrillation
456154015	312442005	456154015	h/o: atrial fibrillation
2675253013	426749004	2675253013	chronic atrial fibrillation
2675306013	427665004	2675306013	paroxysmal atrial flutter
2692063011	428076002	2692063011	history of atrial flutter
636721000000112	440028005	2793259018	permanent atrial fibrillation
636701000000115	440059007	2793372019	persistent atrial fibrillation
329991100000611	6 49436004	1230726010	af - atrial fibrillation
82343012	49436004	82343012	atrial fibrillation
9988012	5370000	9988012	atrial flutter
260821100000011	5 720448006	3320796013	typical atrial flutter
780363100000611	5 720448006	3320801015	atrial flutter type i
780364100000611	3 720448006	3320802010	atrial flutter type 1
175587100000011	7 758600000	3620630016	referral to atrial fibrillation clinic

Snomed codes for Oral Anticoagulants for CPRD Aurum

prodcodeid	productname (termfromemis if productname missing)	drugsub
1337441000033110	sinthrome tablets 4 mg	Acenocoumarol
3097541000033112	acenocoumarol 1mg tablets	Acenocoumarol
1337341000033116	sinthrome 1mg tablets	Acenocoumarol
6444541000033110	eliquis 2.5mg tablets	Apixaban
6444441000033114	apixaban 2.5mg tablets	Apixaban
8232941000033119	apixaban 5mg tablets	Apixaban
8233041000033112	eliquis 5mg tablets	Apixaban
4500341000033115	dabigatran etexilate 110mg capsules	Dabigatran
4500541000033110	pradaxa 110mg capsules	Dabigatran
6436141000033111	dabigatran etexilate 150mg capsules	Dabigatran
6436241000033116	pradaxa 150mg capsules	Dabigatran
4500441000033114	pradaxa 75mg capsules	Dabigatran
4500241000033113	dabigatran etexilate 75mg capsules	Dabigatran
10493841000033112	lixiana 15mg tablets	Edoxaban
10493541000033110	edoxaban 15mg tablets	Edoxaban
10493641000033111	edoxaban 30mg tablets	Edoxaban
10493941000033116	lixiana 30mg tablets	Edoxaban
10494041000033119	lixiana 60mg tablets	Edoxaban
10493741000033119	edoxaban 60mg tablets	Edoxaban
1081541000033110	phenindione 10mg tablets	Phenindione
1082341000033112	phenindione 25mg tablets	Phenindione
1082441000033118	phenindione 50mg tablets	Phenindione
9121441000033118	marcoumar 3mg tablets	Phenprocoumon
9121341000033112	phenprocoumon 3mg tablets	Phenprocoumon
12407141000033118	xarelto 15mg / 20mg treatment initiation pack	Rivaroxaban
12407041000033117	rivaroxaban 15mg tablets and rivaroxaban 20mg tablets	Rivaroxaban
4656441000033114	xarelto 10mg tablets	Rivaroxaban
4656341000033115	rivaroxaban 10mg tablets	Rivaroxaban
6511541000033112	xarelto 15mg tablets	Rivaroxaban
6511341000033117	rivaroxaban 15mg tablets	Rivaroxaban
9704641000033115	xarelto 2.5mg tablets	Rivaroxaban
9704541000033116	rivaroxaban 2.5mg tablets	Rivaroxaban
6511441000033111	rivaroxaban 20mg tablets	Rivaroxaban
6511641000033113	xarelto 20mg tablets	Rivaroxaban
1532041000033119	warfarin wbp tablets 1 mg	Warfarin
1531941000033113	warfarin wbp tablets	Warfarin
1532341000033117	warfarin (evans) tablets 3 mg	Warfarin
1532241000033110	warfarin wbp tablets 5 mg	Warfarin
1532141000033115	warfarin wbp tablets 3 mg	Warfarin
6000741000033110	warfarin 5mg/5ml oral solution	Warfarin
6000541000033119	warfarin 1mg/5ml oral solution	Warfarin
6000641000033118	warfarin 3mg/5ml oral solution	Warfarin
2620141000033114	warfarin 5mg/5ml oral suspension	Warfarin
6066441000033118	warfarin 1mg/ml oral suspension sugar free	Warfarin
2639041000033119	warfarin 10mg/5ml oral suspension	Warfarin
6000841000033117	warfarin 1mg/5ml oral suspension	Warfarin
6000941000033113	warfarin 3mg/5ml oral suspension	Warfarin
1531641000033118	warfarin 1mg tablets	Warfarin
868141000033114	marevan 1mg tablets	Warfarin
1531741000033110	warfarin 3mg tablets	Warfarin
868241000033119	marevan 3mg tablets	Warfarin
6112441000033111	coumadin 4mg tablets	Warfarin
868341000033112	marevan 5mg tablets	Warfarin
1531841000033117	warfarin 5mg tablets	Warfarin
1819641000033114	warfarin 500microgram tablets	Warfarin
1819741000033117	marevan 500microgram tablets	Warfarin

Snomed codes for Stroke TIA SE for CPRD Aurum

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
1263010000061131	1055001	2858018	stenosis of precerebral arteries
2515171000006111	1055001	2861017	narrowing of precerebral artery
1262640100000611 5	111098100000010 4	2779941000000115	qof (quality and outcomes framework) stroke and transient ischaemic attack quality indicator-related care invitation
360778011	111298007	360778011	chronic cerebral ischaemia
4193661000006110	111298007	178529015	chronic cerebral ischemia
966981000006112	112901000000108	199461000000113	ref to multidisciplinary stroke function improvement service
370661000006114	125081000119106	3042974014	[x]cereb infarct due unsp occlus/stenos precerebr arteries
542261000006114	125081000119106	3042974014	cereb infarct due unsp occlus/stenos precerebr arteries
2520541000006112	1386000	3421016	intracranial hemorrhage
300298011	1386000	475553012	intracranial haemorrhage nos
25897016	15258001	25897016	subclavian steal syndrome
1573101000006112	157310100000610 8	1573101000006112	cerebral infarction with haemorrhagic transformation
251692018	161511000	251692018	h/o: tia
4540591000006113	161511000	2986735018	history of transient ischemic attack
1682791000006112	168279100000610 8	1682791000006112	mitoch myopath/encephalopath/lactic acidosis/stroke-like episode
264499015	170600009	264499015	stroke monitoring
1726581000006116	172658100000610 0	1726581000006116	central post-stroke pain
1773371000006117	177337100000610 1	1773371000006117	cerebrovascular accident care plan
1863321000006113	186332100000610 9	1863321000006113	referral by stroke nurse specialist
296940016	192759008	296937016	thrombosis of central nervous system venous sinus nos
4768491000006118	192759008	1784812010	thrombosis of intracranial venous sinus
100721000006111	192759008	296937016	thrombosis of central nervous system venous sinuses
296938014	192760003	296938014	thrombosis of superior longitudinal sinus
296939018	192761004	296939018	thrombosis transverse sinus
4777721000006118	195155004	300243018	subarachnoid hemorrhage from carotid siphon and bifurcation
300244012	195155004	300244012	subarachnoid haemorrhage from carotid siphon and bifurcation
300253017	195160000	300253017	subarachnoid haemorrhage from vertebral artery
4777751000006110	195160000	2915522014	intracranial subarachnoid hemorrhage from vertebral artery
4777761000006112	195160000	2916525016	intracranial subarachnoid haemorrhage from vertebral artery
4777741000006113	195160000	300254011	subarachnoid hemorrhage from vertebral artery
4777811000006117	195165005	300271012	basal ganglia hemorrhage
503791000006114	195165005	300272017	basal nucleus haemorrhage
4777831000006111	195167002	300275015	external capsule hemorrhage
300276019	195167002	300276019	external capsule haemorrhage
4777871000006114	195168007	2916313015	intracerebral hemorrhage with intraventricular hemorrhage

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
477786100000611 9	195168007	2915438018	intracerebral haemorrhage with intraventricular haemorrhage
477785100000611 6	195168007	300278018	intracerebral hemorrhage, intraventricular
300277011	195168007	300277011	intracerebral haemorrhage, intraventricular
746571000006116	195169004	300280012	intracerebral haemorrhage, multiple localized
477788100000611 2	195169004	300280012	intracerebral haemorrhage, multiple localised
477789100000611 0	195169004	300279014	intracerebral hemorrhage, multiple localized
300303013	195180004	300303013	basilar artery occlusion
884461000006113	195180004	884461000006113	basilar artery occluded
477797100000611 5	195180004	2535860015	basilar artery obstruction
477798100000611 7	195180004	2966578019	occlusion of basilar artery
884481000006115	195182007	884481000006115	vertebral artery occluded
300309012	195182007	300309012	vertebral artery occlusion
300310019	195183002	300310019	multiple and bilateral precerebral arterial occlusion
300312010	195185009	300312010	cerebral infarct due to thrombosis of precerebral arteries
300313017	195186005	300313017	cerebral infarction due to embolism of precerebral arteries
300321011	195189003	300321011	cerebral infarction due to thrombosis of cerebral arteries
300322016	195190007	300322016	cerebral infarction due to embolism of cerebral arteries
67501000006115	195199008	300343015	vertebro-basilar artery syndrome
477807100000611 8	195200006	2966553018	carotid artery syndrome
300344014	195200006	300344014	carotid artery syndrome hemispheric
300345010	195201005	300345010	multiple and bilateral precerebral artery syndromes
300352012	195205001	300352012	impending cerebral ischaemia
477810100000611 1	195205001	300351017	impending cerebral ischemia
300353019	195206000	300353019	intermittent cerebral ischaemia
477812100000611 8	195206000	300354013	intermittent cerebral ischemia
300362017	195209007	300362017	middle cerebral artery syndrome
300363010	195210002	300363010	anterior cerebral artery syndrome
300364016	195211003	300364016	posterior cerebral artery syndrome
524511000006116	195212005	300365015	brain stem stroke syndrome
300366019	195213000	300366019	cerebellar stroke syndrome
477818100000611 9	195216008	300369014	left sided cerebral hemisphere cerebrovascular accident
300370010	195216008	300370010	left sided cva
300371014	195217004	300371014	right sided cva
477820100000611 8	195217004	300372019	right sided cerebral hemisphere cerebrovascular accident
586241000006116	195229008	300392014	nonpyogenic venous sinus thrombosis
542251000006112	195230003	300393016	cereb infarct due cerebral venous thrombosis, nonpyogenic
300395011	195232006	300395011	occlusion and stenosis of middle cerebral artery
300396012	195233001	300396012	occlusion and stenosis of anterior cerebral artery
300398013	195234007	300398013	occlusion and stenosis of posterior cerebral artery
300399017	195235008	300399017	occlusion and stenosis of cerebellar arteries
267311000006118	195236009	300400012	occlusion+stenosis of multiple and bilat cerebral arteries

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
300406018	195240000	300406018	sequelae of subarachnoid haemorrhage
4778331000006118	195240000	300405019	sequelae of subarachnoid hemorrhage
300407010	195241001	300407010	sequelae of intracerebral haemorrhage
4778351000006113	195241001	300408017	sequelae of intracerebral hemorrhage
300411016	195243003	300411016	sequelae of cerebral infarction
300533017	195317001	300533017	embolism and thrombosis of the thoracic aorta
884651000006113	195317001	884651000006113	embolus/thrombosis aorta nos
300534011	195318006	300534011	embolism and thrombosis of an arm or leg artery
300550015	195318006	300534011	peripheral arterial embolism and thrombosis nos
300547018	195318006	300534011	embolism and thrombosis of a leg artery nos
300538014	195318006	300534011	embolism and thrombosis of an arm artery nos
300535012	195319003	300535012	embolism and thrombosis of the brachial artery
300536013	195320009	300536013	embolism and thrombosis of the radial artery
300537016	195321008	300537016	embolism and thrombosis of the ulnar artery
300539018	195323006	300539018	embolism and thrombosis of the femoral artery
300540016	195324000	300540016	embolism and thrombosis of the popliteal artery
300541017	195325004	300541017	embolism and thrombosis of the anterior tibial artery
300542012	195326003	300542012	embolism and thrombosis of the dorsalis pedis artery
300543019	195327007	300543019	embolism and thrombosis of the posterior tibial artery
300553018	195335005	300553018	embolism and/or thrombosis of the common iliac artery
300557017	195339004	300557017	embolism and thrombosis of the subclavian artery
300558010	195340002	300558010	embolism and thrombosis of the splenic artery
300559019	195341003	300559019	embolism and thrombosis of the axillary artery
4778981000006116	195342005	300560012	embolism and thrombosis of the celiac artery
300561011	195342005	300561011	embolism and thrombosis of the coeliac artery
4778991000006118	195343000	3491726018	embolism and thrombosis of hepatic artery
300562016	195343000	3491726018	embolism and thrombosis of the hepatic artery
1994761000006112	199476100000610 8	1994761000006112	referral to stroke rehabilitation service
307780013	200258006	307780013	obstetric cerebral venous thrombosis
307781012	200259003	307781012	cerebral venous thrombosis in pregnancy
307782017	200260008	307782017	cerebral venous thrombosis in the puerperium
370701000006118	20059004	33759015	[x]cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs
543141000006110	20059004	33759015	cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs
300943012	20059004	33759015	[x]occlusion and stenosis of other cerebral arteries
2818821000006111	20059004	2966627019	occlusion of cerebral artery
1222398015	20059004	1222398015	cerebral arterial occlusion

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
2008521000006110	200852100000610	2008521000006110	referral to community stroke service
2008601000006116	200860100000610	2008601000006116	discharge from community stroke service
306621000000116	201501000000108	306621000000116	seen in stroke clinic
100701000006118	21258007	1222620011	thrombosis lateral sinus
283799100000611	21258007	35719015	thrombosis of transverse sinus
283798100000611	21258007	35718011	thrombosis of lateral venous sinus
283800100000611	21258007	1222620011	cerebral venous thrombosis of lateral sinus
481028017	21454007	481028017	subarachnoid haemorrhage
428181000006115	21454007	36011016	[x]subarachnoid haemorrh from intracranial artery, unspecif
284115100000611	21454007	1217630014	sah - subarachnoid hemorrhage
284116100000611	21454007	2916299018	subarachnoid intracranial haemorrhage
300935019	21454007	36011016	[x]subarachnoid haemorrhage from other intracranial arteries
284117100000611	21454007	2916568015	subarachnoid intracranial hemorrhage
123481000006118	21454007	36011016	subarachnoid haemorrhage from anterior communicating artery
123521000006118	21454007	36011016	subarachnoid haemorrhage from posterior communicating artery
123511000006114	21454007	36011016	subarachnoid haemorrhage from middle cerebral artery
300257016	21454007	481028017	subarachnoid haemorrhage nos
284114100000611	21454007	1216125018	sah - subarachnoid haemorrhage
300936018	21454007	481028017	[x]other subarachnoid haemorrhage
123441000006112	21454007	36011016	subarachnoid haemorrh from intracranial artery, unspecif
154805100000611	223501000000102	356141000000111	perc translum embolis major systemic pulmonary collater art
405339016	230690007	345637012	stroke and cerebrovascular accident unspecified
605491000006113	230690007	345635016	cva - cerebrovascular accident unspecified
605501000006117	230690007	345637012	cva unspecified
501098100000611	230690007	345636015	stroke
122401000006115	230690007	345637012	stroke unspecified
118785110000061	230691006	3636108019	cerebrovascular accident due to occlusion of cerebral artery
501102100000611	230691006	2914970017	cerebrovascular accident due to cerebral artery occlusion
122361000006113	230691006	345638019	stroke due to cerebral arterial occlusion
605461000006117	230691006	345638019	cva - cerebral artery occlusion
345639010	230692004	345639010	infarction - precerebral
299342019	230698000	345651012	[x]other lacunar syndromes
501116100000611	230698000	345652017	lacunar stroke
501117100000611	230698000	345653010	laci - lacunar infarction
501118100000611	230698000	345654016	li - lacunar infarction
345655015	230699008	345655015	pure motor lacunar syndrome
501119100000611	230699008	345656019	pure motor lacunar infarction
345658018	230700009	345658018	pure sensory lacunar syndrome
501121100000611	230700009	345657011	pure sensory lacunar infarction

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
345675012	230710000	345675012	lobar cerebral haemorrhage
501140100000611 0	230710000	345674011	lobar cerebral hemorrhage
345684012	230716006	345684012	carotid territory transient ischaemic attack
501150100000611 1	230716006	345685013	anterior circulation transient ischemic attack
501151100000611 4	230716006	345686014	carotid territory transient ischemic attack
501149100000611 5	230716006	345683018	anterior circulation transient ischaemic attack
371641000000112	231231000000107	371641000000112	delivery of rehabilitation for stroke
350563016	233972005	350563016	saddle embolus
356328018	237766002	356328018	adrenocortical haemorrhage
511220100000611 1	237766002	356329014	adrenocortical hemorrhage
395777014	266253001	395777014	precerebral arterial occlusion
300942019	266253001	395777014	[x]occlusion and stenosis of other precerebral arteries
300311015	266253001	395777014	other precerebral artery occlusion
122229710000061 19	266253001	3644064018	occlusion of precerebral artery
300314011	266253001	395777014	precerebral artery occlusion nos
395778016	266254007	395778016	carotid artery occlusion
884471000006118	266254007	884471000006118	carotid artery occluded
118232010000061 16	266254007	3636115010	occlusion of carotid artery
395783012	266257000	395783012	transient ischaemic attack
988951000006117	266257000	988951000006117	transient ischaemic attacks
300348012	266257000	395788015	other transient cerebral ischaemia
300349016	266257000	395788015	transient cerebral ischaemia nos
95931000006111	266257000	395788015	transient cerebral ischaemia nos
416991000006112	266257000	395788015	[x]other transnt cerebral ischaemic attacks+related syndroms
549220100000611 1	266257000	395784018	tia - transient ischaemic attack
549218100000611 0	266257000	395782019	temporary cerebral vascular dysfunction
549217100000611 2	266257000	395781014	tia
549222100000611 8	266257000	395787013	transient ischemic attack
395788015	266257000	395788015	transient cerebral ischaemia
119201210000061 17	266257000	395785017	transient cerebral ischemia
395794011	266262004	395794011	arterial embolic and thrombotic occlusion
884631000006118	266262004	884631000006118	arterial embolism/thrombosis
300556014	266262004	395794011	embolism and thrombosis of the iliac artery unspecified
300564015	266262004	395794011	arterial embolism and thrombosis nos
300964011	266262004	395794011	[x]embolism and thrombosis of other arteries
395795012	266262004	395795012	arterial embolus and thrombosis
300563014	266262004	395794011	embolism and thrombosis of other arteries nos
491241000006118	266262004	395795012	arterial embolism and thrombosis
300552011	266262004	395794011	embolism and thrombosis of other specified artery
395796013	266263009	395796013	embolism and thrombosis of the abdominal aorta
884641000006111	266263009	884641000006111	embolus/thrombosis abd. aorta
1227592012	266995000	397829016	[v]personal history of cerebrovascular accident (cva)
989211000006119	274100004	989211000006119	cerebral haemorrhage nos
300287010	274100004	2819959010	intracerebral haemorrhage nos

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
122371000006118	274100004	409859018	stroke due to intracerebral haemorrhage
884421000006119	274100004	884421000006119	cerebral haemorrhage
300939013	274100004	409860011	[x]other intracerebral haemorrhage
744901000006114	274100004	2819959010	intracerebral haemorrhage
744921000006116	274100004	409860011	intracerebral haemorrhage in hemisphere, unspecified
122231110000061 15	274100004	3673216015	intracerebral hemorrhage
119035710000061 10	274100004	122371000006118	stroke due to intracerebral haemorrhage
122231310000061 14	274100004	3673218019	ich - intracerebral hemorrhage
556016100000611 1	274100004	2819960017	intracerebral hemorrhage (ich)
748941000006115	274100004	409859018	left sided intracerebral haemorrhage, unspecified
300956017	274100004	409860011	[x]intracerebral haemorrhage in hemisphere, unspecified
122231010000061 18	274100004	3673215016	intracerebral haemorrhage
122231210000061 11	274100004	3673217012	ich - intracerebral haemorrhage
605471000006112	274100004	409859018	cva - cerebrovascular accid due to intracerebral haemorrhage
119195710000061 10	274100004	411291000006111	[x]other intracerebral haemorrhage
989201000006117	274100004	989201000006117	cerebral haemorrhage
638871000006114	274101000	409861010	embolism and thrombosis of other and unspec parts aorta
556017100000611 6	274101000	409861010	aortic thromboembolism
605481000006110	275434003	411416011	cva - cerebrovascular accident in the puerperium
411416011	275434003	411416011	stroke in the puerperium
411518010	275526006	411518010	h/o: cva
2476091017	275526006	2476091017	h/o: stroke
1227591017	275526006	2476091017	[v]personal history of stroke
809421000006116	275526006	2986886017	h/o: cva/stroke
123491000006115	276284000	412361011	subarachnoid haemorrhage from basilar artery
558310100000611 6	276284000	412361011	subarachnoid haemorrhage from basilar artery aneurysm
558311100000611 8	276284000	412362016	subarachnoid hemorrhage from basilar artery aneurysm
559704100000611 4	277286006	413749011	cpsp - central post-stroke pain
413750011	277286006	413750011	central post-stroke pain
295023100000611 1	28048009	2950231000006111	subarachnoid haemorrhage following injury without open intracranial wound
119237910000061 14	28048009	46957015	subarachnoid hemorrhage following injury without open intracranial wound
123331000006114	28048009	46957015	subarach h'ge inj no open intracran wnd+loc unspec duration
123321000006111	28048009	46957015	subarach h'ge inj no open intracran wnd+>24hrs loc-restored
123411000006113	28048009	46957015	subarachnoid h'ge inj no open intracran wnd+no loss consc
123311000006115	28048009	46957015	subarach h'ge inj no open intracran wnd + concussion unspec
123421000006117	28048009	46957015	subarachnoid h'ge inj no open intracran wound + 1-24hr loc
123391000006113	28048009	46957015	subarachnoid h'ge inj no open intracran wnd+<1hr loss consc
123431000006119	28048009	46957015	subarachnoid h'ge inj no open intracran wound + unspec consc

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
123401000006110	28048009	46957015	subarachnoid h'ge inj no open intracran wnd+>24 loc+recovery
402929011	28048009	483737017	closed traumatic subarachnoid haemorrhage
300380014	29322000	49074018	acute cerebrovascular insufficiency nos
596082100000611 7	306802002	449934011	referral to stroke service
449935012	306803007	449935012	admission to stroke unit
451133011	307766002	451133011	left sided cerebral infarction
451134017	307767006	451134017	right sided cerebral infarction
451371010	308067002	451371010	h/o: stroke in last year
597452100000611 6	308067002	2986393012	history of stroke in last year
163261000006119	308128006	704642018	right sided intracerebral haemorrhage, unspecified
127224810000061 16	308128006	451441015	right sided intracerebral haemorrhage, unspecified
216711000006118	312377009	456079019	post radiological embolism of upper limb artery
601556100000611 2	312377009	456079019	post-radiological embolism of upper limb artery
601560100000611 2	312380005	456083019	post-radiological embolism of lower limb artery
216701000006116	312380005	456083019	post radiological embolism of lower limb artery
2476647018	34781003	2476647018	vertebral artery compression syndrome
58046010	34781003	58046010	vertebral artery syndrome
1212072018	373606000	1212072018	occlusive stroke
841051000006113	390936003	1477210016	cereb autosom dominant arteriop subcort infarcts leukoenceph
651654100000611 9	390936003	1476249012	cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy
271757017	397045002	1776771012	open embolectomy of coeliac artery nec
271780013	397045002	1776771012	open embolectomy of visceral branch of abdominal aorta nec
394347014	397045002	1776771012	percutaneous transluminal embolectomy of artery nec
658099100000611 8	397045002	1776771012	arterial embolectomy
271759019	397045002	1776771012	open embolectomy of superior mesenteric artery nec
272097014	397045002	1776771012	open embolectomy of artery nec
271760012	397045002	1776771012	open embolectomy of inferior mesenteric artery nec
271761011	397045002	1776771012	open embolectomy of suprarenal artery nec
2474330011	412773009	2474330011	referral to stroke clinic
683705100000611 9	413102000	2966612019	basal ganglion stroke
683706100000611 7	413102000	2966650013	basal ganglion infarct
2474651019	413102000	2474651019	infarction of basal ganglia
683738100000611 3	413124000	2694676019	stroke/transient ischemic attack referral
683737100000611 0	413124000	2469839018	stroke / transient ischemic attack referral
2469365010	413124000	2469365010	stroke / transient ischaemic attack referral
2534253014	415628004	2534253014	stroke/transient ischaemic attack monitoring first letter
688044100000611 9	415628004	2534252016	stroke/transient ischemic attack monitoring first letter
2534212018	415629007	2534212018	stroke/transient ischaemic attack monitoring second letter

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
6880461000006115	415629007	2534211013	stroke/transient ischemic attack monitoring second letter
2534196010	415631003	2534196010	stroke/transient ischaemic attack monitoring third letter
6880501000006115	415631003	2534195014	stroke/transient ischemic attack monitoring third letter
9868310000061150	417059002	2549607012	stroke/transient ischaemic attack monitoring verbal invitati
6901471000006110	417059002	2549606015	stroke/transient ischemic attack monitoring verbal invitation
2548554013	417506008	2548554013	haemorrhagic stroke monitoring
6907881000006118	417506008	2548553019	hemorrhagic stroke monitoring
218511000000117	432504007	2770034014	infarction - cerebral
395780010	432504007	2770034014	cerebral infarction nos
300941014	432504007	2770034014	[x]other cerebral infarction
391043019	450375008	2915656010	traumatic subarachnoid haemorrhage
391042012	450375008	2915309017	subarachnoid haemorrhage following injury
7377831000006117	450375008	2915309017	traumatic hemorrhage into subarachnoid space of neuraxis
505324014	450418003	2916058017	traumatic cerebral haemorrhage
505322013	450418003	2916363013	cerebral haemorrhage following injury
7378441000006113	450418003	2916150010	traumatic cerebral hemorrhage
737840100000611	450418003	2915288012	traumatic intracerebral hemorrhage
737841100000614	450418003	2915447014	cerebral hemorrhage following injury
737842100000618	450418003	2915664016	traumatic intracerebral haemorrhage
39701000006110	450418003	2915288012	oth cereb h'ge inj + open intracran wnd+>24hr loc -restored
39801000006119	450418003	2915288012	oth cerebral h'ge inj + open intracranial wnd+no loss consc
35791000006117	450418003	2915288012	other cerebral h'ge after injury + open intracranial wound
39751000006114	450418003	2915288012	oth cereb h'ge inj no open intracran wnd+concussion unspec
39771000006116	450418003	2915288012	oth cerebral h'ge inj + open intracran wnd + unspec consc
30981000006112	450418003	2915288012	oth cerebral h'ge inj no open intracran wnd+unspec consc
39711000006113	450418003	2915288012	oth cereb h'ge inj + open intracran wnd+concussion unspec
39761000006111	450418003	2915288012	oth cereb h'ge inj no open intracran wnd+loc unspec duration
39791000006115	450418003	2915288012	oth cerebral h'ge inj + open intracran wnd+1-24hr loss consc
39831000006110	450418003	2915288012	oth cerebral h'ge inj no open intracranial wnd+no loss consc
39741000006112	450418003	2915288012	oth cereb h'ge inj no open intracran wnd+>24hr loc -restored
39731000006119	450418003	2915288012	oth cereb h'ge inj no open intracran wnd+>24hr loc +recovery
39781000006118	450418003	2915288012	oth cerebral h'ge inj + open intracran wnd+<1hr loss consc
39811000006116	450418003	2915288012	oth cerebral h'ge inj no open intracran wnd+<1hr loss consc
320897016	450418003	2916363013	other cerebral haemorrhage following injury nos
320852013	450418003	2916363013	cerebral haemorrhage following injury nos
39721000006117	450418003	2915288012	oth cereb h'ge inj + open intracran wnd+loc unspec duration

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
35801000006116	450418003	2915288012	other cerebral h'ge after injury no open intracranial wound
39691000006110	450418003	2915288012	oth cereb h'ge inj + open intracran wnd+>24hr loc + recovery
39821000006112	450418003	2915288012	oth cerebral h'ge inj no open intracran wnd+1-24hr loc
320853015	450418003	2916363013	other cerebral haemorrhage following injury
495394013	49422009	495394013	cortical haemorrhage
329962100000611	49422009	82320015	cortical hemorrhage
9			
116047100000011	519751000000106	1160471000000111	stroke 6 month review
1			
819543100000611	519751000000106	1550221000000113	stroke/cerebrovascular accident 6
8			month review
819542100000611	519751000000106	1160481000000113	cerebrovascular accident (cva) 6 month
6			review
334605100000611	52201006	86879010	internal capsule hemorrhage
0			
496232015	52201006	496232015	internal capsule haemorrhage
320735017	5251007	496337012	open traumatic subarachnoid haemorrhage
123351000006119	5251007	9804011	subarachnoid h'ge inj + open intracran wnd+concussion unspec
123361000006117	5251007	9804011	subarachnoid h'ge inj + open intracran wound + unspec consc
123301000006118	5251007	9804011	subarach h'ge inj + open intracran wnd+loc unspec duration
258345100000611	5251007	9804011	subarachnoid hemorrhage following
1			injury with open intracranial wound
423221000006117	56267009	93568017	[x]predominantly cortical dementia
696161000006115	56267009	93568017	multi infarct dementia
294656010	56267009	497559016	arteriosclerotic dementia nos
341425100000611	56267009	497560014	vad - vascular dementia
2			
341423100000611	56267009	497558012	mid - multi-infarct dementia
7			
341426100000611	56267009	2921000019	multi infarct dementia
4			
497559016	56267009	497559016	arteriosclerotic dementia
363791000006112	56267009	497559016	[x]arteriosclerotic dementia
399031000006111	56267009	93568017	[x]multi-infarct dementia
127331510000061	583731000000103	1295391000000110	peripheral arterial embolism and
14			thrombosis nos
884661000006110	583731000000103	884661000006110	peripheral arterial embolism
884671000006115	583761000000108	884671000006115	embolus/thrombus artery nos
127331610000061	583761000000108	1295451000000114	arterial embolism and thrombosis nos
11			
127332210000061	584181000000100	1296321000000119	transient cerebral ischaemia nos
10			
884511000006111	584181000000100	884511000006111	transient ischaemic attacks
300290016	62914000	104563015	other and unspecified intracranial haemorrhage
106394016	64009001	106394016	vertebrobasilar insufficiency
106392017	64009001	106392017	basilar artery syndrome
499739014	64009001	499739014	insufficiency - basilar artery
354078100000611	64009001	106393010	vertebrobasilar arterial insufficiency
7			
67511000006117	64009001	106394016	vertebro-basilar insufficiency
107332010	64586002	107332010	carotid artery stenosis
126531000006111	64586002	107332010	stenosis, carotid artery
355051100000611	64586002	107333017	carotid artery narrowing
5			
218541000000116	65198009	108347015	thrombosis - arterial
127274310000061	682621000000105	1495311000000117	cerebral infarction nos
19			

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
884501000006113	682621000000105	884501000006113	cerebral a. occlusion nos
127276910000061 17	685631000000102	1501361000000113	stroke and cerebrovascular accident unspecified
884531000006117	685631000000102	884531000006117	stroke
884521000006115	685631000000102	884521000006115	stroke/cva - undefined
212618100000011 4	699270006	2983500010	stroke annual review
751944100000611 2	699270006	2983532010	cerebrovascular accident annual review
1780304018	699270006	2983515011	stroke/cva annual review
236331000000111	713771000000100	1565011000000111	stroke/transient ischaemic attack monitoring administration
118689010	71444005	118689010	cerebral thrombosis
366232100000611 7	71444005	118691019	cerebral arterial thrombosis
366231100000611 3	71444005	118690018	thrombosis of cerebral arteries
366233100000611 9	71444005	1233388010	ct - cerebral thrombosis
139091000000115	716021000000109	1568581000000116	excepted from stroke quality indicators: informed dissent
138981000000113	716331000000108	1568881000000119	exception reporting: stroke quality indicators
139081000000117	716581000000101	1569131000000118	excepted from stroke quality indicators: patient unsuitable
156258100000611 4	717241000000108	1569781000000115	stroke/transient ischaemic attack monitoring telephone invte
483988011	732923001	3467313018	bulbar haemorrhage
795127100000611 6	732923001	3467314012	haemorrhage of medulla oblongata
795128100000611 8	732923001	3467313018	hemorrhage of medulla oblongata
3503835018	734298005	3503835018	thromboembolus of internal iliac artery
300554012	734298005	3503835018	embolism and/or thrombosis of the internal iliac artery
300555013	734299002	3503839012	embolism and/or thrombosis of the external iliac artery
3503839012	734299002	3503839012	thromboembolus of external iliac artery
114833100000011 1	736288002	3517033016	transient ischaemic attack clinical management plan
3517034010	736288002	3517034010	transient ischemic attack clinical management plan
502878012	75038005	502878012	cerebellar haemorrhage
371985100000611 6	75038005	124627016	cerebellar hemorrhage
371986100000611 9	75038005	124628014	haemorrhagic cerebellum
371987100000611 4	75038005	124629018	hemorrhagic cerebellum
166774100000011 0	751371000000107	1667741000000110	[v]personal history of transient ischaemic attack
823115100000611 0	751371000000107	1653121000000119	personal history of transient ischaemic attack
125470015	75543006	125470015	cerebral embolism
372845100000611 6	75543006	125471016	cerebral arterial embolism
542831000006116	75543006	125470015	cerebral embolus
503469016	7713009	503469016	pontine haemorrhage
262264100000611 6	7713009	13755017	intrapontine hemorrhage
262263100000611 4	7713009	503468012	intrapontine haemorrhage
262265100000611 9	7713009	13756016	pontine hemorrhage

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
178465100000611	781731000000103	1747751000000110	ref multidisciplinary stroke function improvement declined
377798100000611	78569004	130373013	posterior inferior cerebellar artery syndrome
377802100000611	78569004	1234269010	lms - lateral medullary syndrome
57341000006119	78569004	130374019	wallenberg syndrome
377801100000611	78569004	130376017	inferior cerebellar artery syndrome
130375018	78569004	130375018	lateral medullary syndrome
211582100000011	810991000000109	2115821000000118	stroke self-management plan review
211801100000011	812041000000104	2118011000000114	stroke self-management plan agreed
212778100000011	816561000000108	2127781000000110	stroke initial post discharge review
100771000006112	86003009	142588012	thrombosis, carotid artery
100691000006118	89980009	1235519010	thrombosis cavernous sinus
396258100000611	89980009	149150017	thrombosis of cavernous venous sinus
905541000006119	905541000006103	905541000006119	[rfc] arterial embolism of limbs
907581000006119	907581000006103	907581000006119	[rfc] stroke/cva
907591000006116	907591000006100	907591000006116	[rfc] stroke
908801000006114	908801000006105	908801000006114	[rfc] stroke
909171000006115	909171000006104	909171000006115	[rfc] cva
190065100000611	914991000000106	2350961000000113	scpe class predom patt c.3 infarct of middle cerebral artery
832859100000611	914991000000106	2350991000000119	scpe (surveillance of cerebral palsy in europe) predominant pattern classification c.3 - infarct of the middle cerebral artery
405697100000611	95455008	158110019	thrombosis of cerebral veins
405698100000611	95455008	158111015	cerebral venous thrombosis
1235913017	95455008	1235913017	cerebral vein thrombosis
345650013	95457000	158113017	brainstem infarction nos
524541000006117	95457000	158113017	brainstem infarction
405705100000611	95457000	2966602014	infarction of brain stem
405704100000611	95457000	2966556014	brain stem stroke
405703100000611	95457000	158114011	brain stem infarct
158118014	95460007	158118014	cerebellar infarction
243673100000011	955491000000106	2436731000000112	old cerebral infarction on imaging

CHF LVEF Snomed Codelist

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
18472010	10633002	18472010	acute congestive heart failure
126236110000061 14	110871100000010 8	2774561000000111	excepted from heart failure quality indicators - service unavailable
126263510000061 11	111093100000010 3	2779841000000112	qof (quality and outcomes framework) heart failure quality indicator-related care invitation
206703015	128404006	206703015	right heart failure
216184014	134378009	216184014	congestive heart failure monitoring
216207010	134401001	216207010	left ventricular systolic dysfunction
216246012	134440006	216246012	referral to heart failure clinic
153938100000611 7	153938100000610 1	1539381000006117	heart failure lifestyle plan commenced
153939100000611 9	153939100000610 3	1539391000006119	heart failure information starter pack provided
153947100000611 1	153947100000610 7	1539471000006111	heart failure monitoring - unstable symptoms
153948100000611 4	153948100000610 5	1539481000006114	heart failure monitoring - specialist clinical needs
153949100000611 2	153949100000610 8	1539491000006112	heart failure monitoring - social issues
153950100000611 6	153950100000610 0	1539501000006116	heart failure monitoring - psychological issues
153951100000611 8	153951100000610 2	1539511000006118	heart failure monitoring - multiple readmissions
153952100000611 4	153952100000610 5	1539521000006114	heart failure monitoring - co-medications
153953100000611 2	153953100000610 8	1539531000006112	heart failure monitoring - co-morbidities
153954100000611 9	153954100000610 3	1539541000006119	heart failure monitoring - palliative care
157632100000611 3	157632100000610 9	1576321000006113	cause of death- congestive cardiac failure
454052100000611 1	161505003	2986453013	history of heart failure
251680018	161505003	251680018	h/o: heart failure
174769100000611 9	174769100000610 3	1747691000006119	emergency heart failure admission since last appointment
182216100000611 6	182216100000610 0	1822161000006116	heart failure pathway protocol not followed
182409100000611 9	182409100000610 3	1824091000006119	heart failure clinical pathway protocol followed
185636100000611 6	185636100000610 0	1856361000006116	heart failure monitoring in primary care
185637100000611 1	185637100000610 7	1856371000006111	heart failure monitoring in secondary care
185638100000611 4	185638100000610 5	1856381000006114	heart failure monitoring default
186173100000611 4	186173100000610 5	1861731000006114	auras-af - consider the patient to have heart failure
301694014	19242006	479262018	pulmonary oedema nos
280534100000611 9	19242006	32441014	pulmonary edema
504901000006118	194767001	299653017	benign hypertensive heart disease with ccf
741701000006114	194779001	299672017	hypertensive heart&renal dis wth (congestive) heart failure
789941000006117	194781004	299674016	hyperten heart&renal dis+both(congestv)heart and renal fail
300179017	195111005	300179017	decompensated cardiac failure
300180019	195112003	300180019	compensated cardiac failure
300190010	195114002	300190010	acute left ventricular failure
300214014	195130005	300214014	post cardiac operation functional disturbance

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
300217019	195130005	300214014	post cardiac operation heart failure nos
1991651000006115	199165100000610 4	1991651000006115	severe left ventricular systolic dysfunction
303361000000111	200171000000102	303361000000111	referred by heart failure nurse specialist
303861000000118	200361000000106	303861000000118	did not attend practice nurse heart failure clinic
308261000000111	202231000000106	308261000000111	heart failure review completed
311561000000117	203791000000106	311561000000117	referred to heart failure education group
316833010	206586007	316833010	congenital cardiac failure
350484012	233924009	350484012	heart failure as a complication of care
404741000000119	247361000000100	404741000000119	heart failure 6 month review
407181000000116	248571000000104	407181000000116	did not attend heart failure clinic
529272100000611 3	250908004	1224530010	lvef - left ventricular ejection fraction
374013010	250908004	374013010	left ventricular ejection fraction
403107019	269299003	403107019	cardiac insufficiency as a complication of care
300884014	274096000	409855012	[x]other specified pulmonary heart diseases
556007100000611 7	274096000	409855012	pulmonary heart disease
884181000006110	274096000	884181000006110	pulmonary heart disease
412678013	276514007	412678013	neonatal cardiac failure
558637100000611 9	276514007	412677015	cardiac failure developing in the perinatal period
538601000006110	302213007	443790017	carotico-cavernous sinus fistula
590427100000611 0	302213007	2646841017	ccf - carotid cavernous fistula
590425100000611 7	302213007	2642659018	carotid cavernous fistula
451426015	308118002	451426015	cardiac failure therapy
453099015	309634009	453099015	h/o: heart failure in last year
599097100000611 3	309634009	2986867013	history of heart failure in last year
6978012	3545003	6978012	diastolic dysfunction
490972013	367363000	490972013	right ventricular failure
181610100000611 3	367363000	490972013	right ventricular failure
1484917012	390884006	1484917012	heart failure follow-up
1484918019	390885007	1484918019	heart failure annual review
1488804017	395105005	1488804017	heart failure confirmed
1489358014	395704004	1489358014	left ventricular diastolic dysfunction
301741013	40541001	492666016	acute pulmonary oedema unspecified
315273100000611 3	40541001	1490485013	pulmonary edema - acute
1490256017	40541001	492666016	pulmonary oedema - acute
315270100000611 7	40541001	67601010	acute edema of lung
315269100000611 7	40541001	67598017	acute pulmonary edema
315271100000611 9	40541001	492667013	acute oedema of lung
1216090015	40541001	67598017	acute oedema of lung, unspecified
315272100000611 0	40541001	1490256017	pulmonary oedema - acute
301743011	40541001	492666016	acute pulmonary oedema nos
2159197017	407596008	2159197017	echocardiogram shows left ventricular systolic dysfunction
2159198010	407597004	2159198010	echocardiogram shows left ventricular diastolic dysfunction
2533628012	414586001	2533628012	left ventricular dysfunction monitoring first letter

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
2533629016	414588000	2533629016	left ventricular dysfunction monitoring second letter
2533630014	414589008	2533630014	left ventricular dysfunction monitoring third letter
2548656015	416158002	2548656015	right ventricular systolic dysfunction
2549089012	416573008	2549089012	left ventricular dysfunction monitoring verbal invite
2549128016	416610007	2549128016	right ventricular diastolic dysfunction
6895651000006113	416683003	6895651000006113	emergency hospital admission for heart failure
2549208013	416683003	3082850014	admit heart failure emergency
12490061000006118	416683003	2549208013	admit heart failure emergency
2549243014	416717003	2549243014	seen in heart failure clinic
2549697018	417146007	2549697018	referral to heart failure nurse
2548316014	417359009	2548316014	seen by community heart failure nurse
2616470012	420300004	2616470012	new york heart association classification - class i
2616472016	420913000	2616472016	new york heart association classification - class iii
2616471011	421704003	2616471011	new york heart association classification - class ii
2616473014	422293003	2616473014	new york heart association classification - class iv
31825410000061117	42343007	493287011	congestive cardiac failure
31825310000061110	42343007	70654011	congestive heart disease
31825510000061115	42343007	493288018	ccf - congestive cardiac failure
493287011	42343007	70653017	congestive cardiac failure
31825610000061118	42343007	493289014	chf - congestive heart failure
70653017	42343007	70653017	congestive heart failure
2645623019	423475008	2645623019	heart failure education
2675255018	426611007	2675255018	congestive heart failure due to valvular disease
2694523019	429589006	2694523019	left ventricular cardiac dysfunction
72934016	43736008	72934016	rheumatic left ventricular failure
164770100000011118	446221000	2883808011	heart failure with normal ejection fraction
732112100000611119	446221000	3496968011	heart failure with preserved ejection fraction
166137100000011112	446221000	1713091000000115	hfnef - heart failure with normal ejection fraction
222750100000011110	446221000	2227501000000110	heart failure with preserved ejection fraction
325734100000611616	46847001	78084015	chronic pulmonary edema
494669012	46847001	494669012	chronic pulmonary oedema
82584011	49584005	82584011	acute cor pulmonale
94251011	56675007	94251011	acute heart failure
305601017	609507007	2966901017	cardiac failure following abortive pregnancy
7499321000006117	609507007	2966901017	induced termination of pregnancy complicated by cardiac failure
884201000006111	639401000000103	884201000006111	pulmonary heart disease nos
12726451000006111	639401000000103	1408271000000114	other chronic pulmonary heart disease
741681000006111	64715009	107545013	hypertensive heart disease nos with ccf
111625010	67189007	111625010	acute pulmonary heart disease
299848013	67189007	111625010	acute pulmonary heart disease nos
7586341000006114	704095000	3012159015	referral to heart failure exercise program

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
308041000000118	704095000	3011242019	referral to heart failure exercise programme
173416100000011	704096004	3011612011	referral to heart failure exercise programme not indicated
758636100000611	704096004	3011502011	referral to heart failure exercise program not indicated
173408100000011	704097008	3011446018	referral to heart failure exercise programme declined
758638100000611	704097008	3010562018	referral to heart failure exercise program declined
226181000000110	713781000000103	1565021000000117	left ventricular dysfunction monitoring administration
308231000000118	713791000000101	1565031000000115	heart failure monitoring administration
407441000000115	715951000000107	1568511000000111	exception reporting: heart failure quality indicators
407001000000113	716411000000109	1568961000000117	left ventricular dysfunction monitoring telephone invite
407101000000114	716621000000101	1569171000000116	heart failure monitoring third letter
407061000000112	716971000000109	1569521000000114	heart failure monitoring first letter
407081000000115	717191000000108	1569731000000119	heart failure monitoring second letter
156194100000611	717481000000104	1570021000000117	excepted heart failure quality indicators: patient unsuitabl
156195100000611	717491000000102	1570031000000115	excepted heart failure quality indicators: informed dissent
407041000000111	717501000000108	1570041000000112	heart failure monitoring verbal invite
406801000000118	717531000000102	1570071000000118	heart failure monitoring telephone invite
119984210000061	762994006	3637505016	nyha (new york heart association) classification class
833381000006119	762994006	3637504017	new york heart assoc classification heart failure symptoms
169391100000011	763641000000102	1693911000000115	referral to heart failure education group declined
170362100000011	764851000000102	1703621000000112	worsening pulmonary oedema
174617100000011	781051000000108	1746171000000119	has heart failure management plan
178406100000611	789621000000105	1765351000000114	preferred place of care for next exacerbation heart failure
132655012	79955004	132655012	chronic cor pulmonale
380029100000611	79955004	1234420011	cor - chronic cor pulmonale
884211000006114	79955004	884211000006114	other pulmonary heart disease
211578100000011	810971000000105	2115781000000114	heart failure self-management plan review
211793100000011	812001000000102	2117931000000116	heart failure self-management plan agreed
212219100000011	813991000000101	2122191000000110	education about deteriorating heart failure
385250100000611	83105008	137848017	malignant hypertensive heart disease with congestive heart failure
728671000006119	83105008	1236017010	malignant hypertensive heart disease with ccf
139482012	84114007	139482012	cardiac failure
139475013	84114007	139475013	heart failure
395772015	84114007	139475013	heart failure nos
386831100000611	84114007	139480016	myocardial failure
386832100000611	84114007	139481017	weak heart
386834100000611	84114007	1234906013	hf - heart failure
223981000000118	84114007	139482012	cardiac failure nos
386835100000611	84114007	2969213019	cardiac insufficiency

medcodeid	snomedctconceptid	SnomedCTDescriptionId	term
139481017	84114007	139475013	weak heart
220595100000011	851521000000102	2205951000000117	heart failure clinical pathway
141306010	85232009	141306010	left ventricular failure
388606100000611	85232009	201199018	left-sided heart failure
388607100000611	85232009	1235017018	lvf - left ventricular failure
388604100000611	85232009	141303019	left heart failure
308301000000118	872361000000105	2253411000000117	heart failure care plan discussed with patient
225681100000011	873881000000100	2256811000000114	referral to rapid access heart failure clinic
299851018	87837008	145620019	other chronic pulmonary heart disease
299853015	87837008	145620019	other chronic pulmonary heart disease nos
395756011	87837008	145620019	chronic pulmonary heart disease nos
884191000006113	87837008	884191000006113	chronic pulmonary heart dis.
145620019	87837008	145620019	chronic pulmonary heart disease
392752100000611	87837008	145622010	chronic cardiopulmonary disease
147247018	88805009	147247018	chronic congestive heart failure
905391000006119	905391000006103	905391000006119	[rfc] cardiac failure
235239100000011	915571000000102	2352391000000117	on optimal heart failure therapy
400530100000611	92506005	510016018	biventricular failure
510016018	92506005	153058012	biventricular failure
939571000006115	939571000006104	939571000006115	diastolic dysfunction
240587100000011	939881000000105	2405871000000117	heart failure rehabilitation programme not available
308011000000119	961881000000101	2451451000000116	heart failure information given to patient

Stroke/TIA/SE HES

ICD-10 Code	term	type
H34.1	central retina artery occlusion	stroke
I63	cerebral infarction	stroke
I64	stroke, not specified as hemorrhage or infarction	stroke
161	intrecerebral hemorrhage	stroke
160	subarachnoid hemorrhage	stroke
I62	other and unspecified nontraumatic intracranial hemorrhage	stroke
G45	transient cerebral ischemia attacks and related syndromes	TIA
I74	arterial embolism and thrombosis	SE
I65	occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction	SE
I66	occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction	SE
I24.0	Acute coronary thrombosis not resulting in myocardial infarction	SE

Stroke Embolism HES

ICD-10 Code	term
I63	cerebral infarction
I64	stroke, not specified as hemorrhage or infarction
161	intrecerebral hemorrhage
160	subarachnoid hemorrhage
I62	other and unspecified nontraumatic intracranial hemorrhage
I74	arterial embolism and thrombosis

Major Bleed HES

ICD-10 Code	term	category
I60	Subarachnoid	Intracranial
I61	intracerebral	Intracranial
I62.0	subdural	Intracranial
I62.1	nontraumatic extradural	Intracranial
I62.9	intracranial, nontraumatic, unspecified	Intracranial
S06.6	Traumatic subarachnoid hemorrhage,	Intracranial
S06.5	Traumatic subdural hemorrhage,	Intracranial
S06.4	Epidural hemorrhage	Intracranial
S06.36	Traumatic hemorrhage of cerebrum, unspecified	Intracranial
S06.34	Traumatic hemorrhage of right cerebrum	Intracranial
S06.35	Traumatic hemorrhage of left cerebrum	Intracranial
S06.36	Traumatic hemorrhage of cerebrum, unspecified	Intracranial
S06.37	Contusion, laceration, and hemorrhage of cerebellum	Intracranial
S06.38	Contusion, laceration, and hemorrhage of brainstem	Intracranial
K92.0	haematemesis	Gastrointestinal
K92.1	melaena	Gastrointestinal
I85.0	oesophageal varices with bleeding	Gastrointestinal
I85.11	Secondary esophageal varices with bleeding	Gastrointestinal
K22.8	Other specified diseases of esophagus (approx. synonym esophageal bleeding)	Gastrointest
K22.11	Ulcer of esophagus with bleeding	Gastrointestinal
K22.6	Gastro-esophageal laceration-hemorrhage syndrome	Gastrointestinal
K31.811	Angiodysplasia of stomach and duodenum with bleeding	Gastrointestinal
K29.21	Alcoholic gastritis with bleeding	Gastrointestinal
K29.31	Chronic superficial gastritis with bleeding	Gastrointestinal
K29.41	Chronic atrophic gastritis with bleeding	Gastrointestinal
K29.51	Unspecified chronic gastritis with bleeding	Gastrointestinal
K29.61	Other gastritis with bleeding	Gastrointestinal
K29.71	Gastritis, unspecified, with bleeding	Gastrointestinal
K29.81	Duodenitis with bleeding	Gastrointestinal
K29.91	Gastroduodenitis, unspecified, with bleeding	Gastrointestinal
I98.20	oesophageal varices in diseases classified elsewhere with bleeding	Gastrointestinal
I98.3	Oesophageal varices with bleeding in disease classified elsewhere	Gastrointestinal
K22.10	Ulcer of oesophagus, acute with bleeding	Gastrointestinal
K22.12	Ulcer of oesophagus, acute with both bleeding and perforation	Gastrointestinal
K22.14	Ulcer of oesophagus, chronic or unspecified with bleeding	Gastrointestinal
K22.16	Ulcer of oesophagus, chronic or unspecified with both bleedingand perforation	Gastrointestinal
K25.0	Gastric ulcer, acute with bleeding	Gastrointestinal
K25.2	Gastric ulcer, acute with both bleeding and perforation	Gastrointestinal
K25.4	Gastric ulcer, chronic or unspecified with bleeding	Gastrointestinal
K25.6	Gastric ulcer, chronic or unspecified with both bleeding and perforation	Gastrointest
K26.0	Duodenal ulcer, acute with bleeding	Gastrointestinal
K26.2	Duodenal ulcer, acute with both bleeding and perforation	Gastrointestinal
K26.4	Duodenal ulcer, chronic or unspecified with bleeding	Gastrointestinal
K26.6	Duodenal ulcer, chronic or unspecified with both bleedingand perforation	Gastrointest
K27.0	Peptic ulcer, acute with bleeding	Gastrointestinal
K27.2	Peptic ulcer, acute with both bleeding and perforation	Gastrointestinal
K27.4	Peptic ulcer, chronic or unspecified with bleeding	Gastrointestinal
K27.6	Peptic ulcer, chronic or unspecified with both bleeding and perforation	Gastrointest
K28.0	Gastrojejunal ulcer, acute with bleeding	Gastrointestinal
K28.2	Gastrojejunal ulcer, acute with both bleeding and perforation	Gastrointestinal
K28.4	Gastrojejunal ulcer, chronic or unspecified with bleeding	Gastrointestinal

K28.6 Gastrojejunal ulcer, chronic or unspecified with both bleeding and perforation Gastrointestinal
K29.0 Acute bleeding gastritis Gastrointestinal
K63.80 Angiodysplasia of small intestine, except duodenum with bleeding Gastrointestinal
K31.80 Angiodysplasia of stomach and duodenum with bleeding Gastrointestinal
K55.21 Angiodysplasia of colon with bleeding Gastrointestinal
K62.5 bleeding of anus and rectum Gastrointestinal
K92.2 Gastrointestinal bleeding, unspecified Gastrointestinal
K57.11 Diverticulosis of small intestine without perforation or abscess with bleeding Gastrointestinal
K57.13 Diverticulitis of small intestine without perforation or abscess with bleeding Gastrointestinal
K57.31 Diverticulosis of large intestine without perforation or abscess with bleeding Gastrointestinal
K57.33 Diverticulitis of large intestine without perforation or abscess with bleeding Gastrointestinal
N02.0 Recurrent and persistent haematuria, minor glomerular abnormality Other
N02.1 Recurrent and persistent haematuria, focal and segmental glomerular lesions Other
N02.2 Recurrent and persistent haematuria, diffuse membranous glomerulonephritis Other
N02.3 Recurrent and persistent haematuria, diffuse mesangial proliferative glomerulonephritis Other
N02.4 Recurrent and persistent haematuria, diffuse endocapillary proliferative glomerulonephritis Other
N02.5 Recurrent and persistent haematuria, diffuse mesangiocapillary glomerulonephritis Other
N02.6 Recurrent and persistent haematuria, dense deposit disease Other
N02.7 Recurrent and persistent haematuria, diffuse crescentic glomerulonephritis Other
N02.8 Recurrent and persistent haematuria, other Other
N02.9 Recurrent and persistent haematuria, unspecified Other
R31.0 Gross hematuria Other
R31.1 Microscopic hematuria Other
R31.8 Other and unspecified hematuria Other
K66.1 Haemoperitoneum Other
N93.8 Other specified abnormal uterine and vaginal bleeding Other
N93.9 Abnormal uterine and vaginal bleeding, unspecified Other
N95.0 Postmenopausal bleeding Other
R04.1 bleeding from throat Other
R04.2 Haemoptysis Other
R04.8 bleeding from other sites in respiratory passages Other
R04.9 bleeding from respiratory passages, unspecified Other
R58 bleeding, not elsewhere classified Other
D68.3 Haemorrhagic disorder due to circulating anticoagulants Other
H35.6 Retinal bleeding Other
H43.1 Vitreous bleeding Other
H45.0 Vitreous bleeding in diseases classified elsewhere Other
H31.3 Choroidal hemorrhage and rupture Other
H21.0 Hyphema Other
H47.02 Hemorrhage in optic nerve sheath Other
M25.0 Haemarthrosis Other
H44.81 Hemophthalmos Other
D62 Acute posthemorrhagic anemia Other
I31.2 Hemopericardium Other
J94.2 Hemothorax Other

CHF LVEF HES

ICD-10 Code	term
I50	Heart failure
I50.1	Left ventricular failure, unspecified
I50.2	Systolic (congestive) heart failure
I50.20	Unspecified systolic (congestive) heart failure
I50.21	Acute systolic (congestive) heart failure
I50.22	Chronic systolic (congestive) heart failure
I50.23	Acute on chronic systolic (congestive) heart failure

I50.3 Diastolic (congestive) heart failure
I50.30 Unspecified diastolic (congestive) heart failure
I50.31 Acute diastolic (congestive) heart failure
I50.32 Chronic diastolic (congestive) heart failure
I50.33 Acute on chronic diastolic (congestive) heart failure
I50.4 Combined systolic (congestive) and diastolic (congestive) heart failure
I50.40 Unspecified combined systolic (congestive) and diastolic (congestive) heart failure
I50.41 Acute combined systolic (congestive) and diastolic (congestive) heart failure
I50.42 Chronic combined systolic (congestive) and diastolic (congestive) heart failure
I50.43 Acute on chronic combined systolic (congestive) and diastolic (congestive) heart failure
I50.8 Other heart failure
I50.81 Right heart failure
I50.810 â€¦â€¦ unspecified
I50.811 Acute right heart failure
I50.812 Chronic right heart failure
I50.813 Acute on chronic right heart failure
I50.814 â€¦â€¦ due to left heart failure
I50.82 Biventricular heart failure
I50.83 High output heart failure
I50.84 End stage heart failure
I50.89 Other heart failure
I50.9 Heart failure, unspecified
I11.0 Hypertensive heart disease with heart failure
I13.0 Hypertensive heart and chronic kidney disease with heart failure and stage 1 through stage 4 chronic kidney disease, or unspecified chronic kidney disease
I13.2 Hypertensive heart and chronic kidney disease with heart failure and with stage 5 chronic kidney disease, or end stage renal disease
I97.13 Postprocedural heart failure
I09.81 Rheumatic heart failure

A2.3 Additional information on methods

A2.3.1 Algorithms

A2.3.1.1 Classification of ethnicity

Ethnicity was self-reported and recorded in CPRD Aurum alone for some patients, in HES alone for other patients, and in both data sources for some patients. Ethnicity was recorded at a country level (such as ‘Polish’ or ‘Somalian’) for some patients and for other patients only broader categories such as ‘Asian’ or ‘Black’ were recorded. To classify patients in consistent categories for ethnicity, a 2-step system was used:

> Step 1: Map ethnicity recorded to the pre-specified categories: White, Black, South Asian, East Asian, Mixed, Other, Unknown
>
> Step 2: Derive ethnicity for each patient by checking both CPRD Aurum and HES and using the following rules to assign ethnicity

For both CPRD Aurum and HES ethnicity:
IF only 1 ethnicity recorded or multiple records all recording the same ethnicity THEN ethnicity = this recorded ethnicity

IF multiple different ethnicities recorded THEN DO
  Compare frequency of ethnicity recorded and select category with the highest count
  IF multiple ethnicities with equivalent highest count THEN DO
    IF different categories are specified categories THEN select most recently recorded
    IF equal counts of 'Other' and a specified ethnicity THEN use the specified ethnicity
IF equal counts of 2 different ethnicities and both recorded on same day (last recorded) then assign ethnicity as 'Mixed'
HES only has 1 record of ethnicity for each patient.

To combine the CPRD Aurum and HES recorded ethnicities the following algorithm was used:
IF only 1 of (CPRD Aurum, HES) ethnicity recorded then use this.
IF both (CPRD Aurum, HES) ethnicity recorded then select CPRD Aurum ethnicity.

A2.3.1.2 Classification of smoking status

Smoking status was recorded in CPRD Aurum in multiple ways and at multiple timepoints. For example, a GP or nurse may record a patient as being a ‘non-smoker’ or ‘smoker’ or may record number of cigarettes consumed per day.

A patient may change smoking status over time meaning looking at data on smoking status over time provides a more accurate picture.

Patients were classified as being 1 of: non-smoker, ex-smoker, or current smoker at their index date.

Smoking records were first mapped to the categories of non-smoker/ex-smoker/current smoker.

Where a patient had only 1 smoking category or all categories the same prior to index date then this category was used as their smoking status.
Where a patient had multiple differing smoking categories prior to index date the following algorithm was followed to assign smoking status:

IF only 1 smoking record prior to index date THEN status = smoking record category
IF multiple smoking records prior to index date all of identical category THEN status = smoking records category
IF multiple smoking records prior to index date of different categories THEN DO
    IF most recent record current-smoker THEN status = current-smoker
    IF most recent record ex-smoker THEN status = ex-smoker
    IF most recent record non-smoker and have earlier records of current or ex-smoker
        THEN status = ex-smoker
END

If a patient had multiple conflicting smoking records on the same day, then the worst option was used in derivation ie current-smoker selected over ex-smoker or non-smoker and ex-smoker selected over non-smoker.

A2.3.1.3 Classification of alcohol consumption

Alcohol status is frequently recorded by GPs most commonly following the GP asking a patient the number of units of alcohol they drink in an average week. A patient's alcohol consumption may be recorded in a descriptive way such as 'tee-total' or 'heavy drinker' or may be inferred by codes used to record referral of a patient to alcohol reduction advice or services. Patient data on alcohol consumption was used to classify patients in the following way:

- Non-drinker (tee total)
- Light drinker (1 to 14 units consumed in an average week)
- Moderate drinker (15 to 42 units consumed in an average week)
- Heavy drinker (>42 units consumed in an average week)

The last record on alcohol consumption recorded prior to index date was used to classify patients. Where patients had conflicting data recorded on alcohol consumption on the same day the heaviest consumption classification was selected.

A2.3.2 Additional information on selection of subset matching ARISTOTLE

A2.3.2.1 Additional information on stroke risk score in ARISTOTLE

The $CHADS_2$ score was the dominant stroke risk score used for patients with AF at the time ARISTOTLE was conducted. This score was superseded by the CHA_2DS_2-VASc score proposed by Lip et al (157), which was found to more accurately predict stroke risk. The additional stroke risk factors from the CHA_2DS_2-VASc score, namely vascular disease defined by peripheral artery disease, MI, or aortic plaque, were therefore included in the propensity score model. $CHADS_2$ is derived by assigning points for different stroke risk factors and taking the sum to obtain a total score (Table A2.3.2.1.1).

Table A2.3.2.1.1 Derivation of $CHADS_2$ Score

$CHADS_2$ Stroke Risk Factor	Points
Congestive Heart Failure (C)[a]	1
Hypertension (H)	1
Age (A) 75 years or older	1
Diabetes (D)	1
Stroke (S) any history of stroke or TIA [b]	2

[a] original definition only included congestive heart failure diagnosis and excluded reduced left ventricle ejection fraction
[b] original definition excluded systemic embolism

The total $CHADS_2$ score can therefore range from 0 to a maximum of 6 with each score corresponding to different possible combinations of risk factors as detailed in Table A2.3.2.1.2. Given the publications of ARISTOTLE summarised the proportion of participants with each $CHADS_2$ score, with each stroke risk factor, in different age categories, male and female, and $CHADS_2$ and stroke risk factor proportions by sex, this information could be combined to construct simultaneous equations that described the different combinations of patient characteristics.

Table A2.3.2.1.2 Stroke Risk Factor Combinations for Each $CHADS_2$ Score and Expected Stroke Rate

$CHADS_2$ Score	**Stroke Risk Factor Combinations**	**Adjusted Stroke Rate % (95% CI)[a]**
0	reduced LVEF; history of SE	1.9 (1.2-3.0)
1	C; H; A; D	2.8 (2.0-3.8)
2	CH; CA; CD; HA; HD; AD; S	4.0 (3.1-5.1)
3	CHA; CHD; CAD; CS; HAD; HS; AS; DS	5.9 (4.6-7.3)
4	CHAD; CHS; CAS; CDS; HAS; HDS; ADS	8.5 (6.3-11.1)
5	CHAS; CHDS; CADS; HADS	12.5 (8.2-17.5)
6	CHADS	18.2 (10.5-27.4)

LVEF = left ventricle ejection fraction; SE = systemic embolism; C = congestive heart failure; H = hypertension; A = age 75 or older; D = diabetes; S = history of stroke or TIA
[a] Adjusted stroke rate is expected stroke rate per 100-patients years taken from 'Validation of Clinical Classification Schemes for Predicting Stroke Results From the National Registry of Atrial Fibrillation' Gage BF et al(61)

A2.3.2.2 Derivation of simultaneous equations describing combinations of characteristics

From the trial publication presented in "Table 1 Baseline Characteristics of the Patients" the combinations of characteristics for sex, prior VKA experience, and number of participants with each stroke risk factor in the apixaban arm were extracted:

$$3234\ women + 5886\ men = 9120$$

$$5208\ VKA_{exp} + 3912\ VKA_{naive} = 9120$$

$$6270\ age_{<75} + 2850\ age_{\geq 75} = 9120$$

Congestive heart failure = C = 2784

Hypertension = H = 7962

Age ≥ 75 years = $A = 2850$

Diabetes = D = 2284

History of stroke or TIA = S approx. 1650

The FDA clinical review of the NDA for apixaban [N Beasley and M Rose, Table 30 (131)] gave the distribution of each $CHADS_2$ score (whereas ARISTOTLE publication grouped $CHADS_2$ 3 or greater together) and age groups in the apixaban arm. Combining these numbers and considering the combinations of stroke risk factors provided initial numbers that could guide potential solutions illustrated in Table A2.3.2.2.1.

Table A2.3.2.2.1 $CHADS_2$ Score by Age-group

	Age < 75		**Age ≥ 75**		**Total**
$CHADS_2$ Score	**(N=6270)**	**No of groups**	**(N=2850)**	**No of groups**	**No of groups**
0 (N=54)	reduced LVEF and /or history of SE	1	N/A	0	1
1 (N=3046)	C; H; D	3	A	1	4
2 (N=3262)	CH; CD; HD; S	4	CA; HA; AD	3	7
3 (N=1681)	CHD; CS; HS; DS	4	CHA; CAD; HAD; AS	4	8
4 (N=767)	CHS; CDS; HDS	3	CHAD; CAS; HAS; ADS	4	7
5 (N=273)	CHDS	1	CHAS; CADS; HADS	3	4
6 (N=37)	N/A	0	CHADS	1	1
Total		16		16	32

LVEF = left ventricle ejection fraction; SE = systemic embolism; N/A = not applicable;
C = congestive heart failure; H = hypertension; A = age 75 or older; D = diabetes; S = history of stroke or TIA;
Combinations of letters represents combinations of risk factors for example HS represents a person with hypertension AND prior stroke.
Target numbers (N=XX) derived from tabulations of baseline characteristics of ARISTOTLE participants.

Referring to Table A2.3.2.2.1 we see equations can be derived relating the number in each $CHADS_2$ score group to the corresponding combinations of $CHADS_2$ stroke risk factors that could make up such a group. If we let $x_{i,j}$ denote the number of participants with $CHADS_2$ score i and stroke risk factor combination j then we have

$$54\ ch_0 = x_{0,1.}\ \textit{(reduced left ventricular ejection fraction [LVEF] and/or SE)}$$

$$3046\ ch_1 = x_{1,1}\,C + x_{1,2}\,H + x_{1,3}\,D + x_{1,4}\,A$$

$$3262\ ch_2 = x_{2,1}\,CH + x_{2,2}\,CD + x_{2,3}\,HD + x_{2,4}\,S + x_{2,5}\,CA + x_{2,6}\,HA + x_{2,7}\,AD$$

$$1681\ ch_3 = x_{3,1}\,CHD + x_{3,2}\,CS + x_{3,3}\,HS + x_{3,4}\,DS + x_{3,5}\,CHA + x_{3,6}\,CAD + x_{3,7}\,HAD + x_{3,8}\,ADS$$

$$767\ ch_4 = x_{4,1}\,CHS + x_{4,2}\,CDS + x_{4,3}\,HDS + x_{4,4}\,CHAD + x_{4,5}\,CAS + x_{4,6}\,HAS + x_{4,7}\,ADS$$

$$273\ ch_5 = x_{5,1}\,CHDS + x_{5,2}\,CHAS + x_{5,3}\,CADS + x_{5,4}\,HADS$$

$$37\ ch_6 = x_{6,1}\,CHADS$$

These 7 equations (for ch_0 through to ch_6) are complimented by an additional 5 equations relating the total number of participants with each $CHADS_2$ risk factor with the combinations that contribute to them; writing these out more clearly by omitting the labels for the subgroups and substituting the value for $x_{6,1}$ into the equations we obtain:

$$2747 = x_{1,1} + x_{2,1} + x_{2,2} + x_{2,5} + x_{3,1} + x_{3,2} + x_{3,5} + x_{3,6} + x_{4,1} + x_{4,2} + x_{4,4} + x_{4,5} + x_{5,1} + x_{5,2} + x_{5,3}$$

$$7925 = x_{1,2} + x_{2,1} + x_{2,3} + x_{2,6} + x_{3,1} + x_{3,3} + x_{3,5} + x_{3,7} + x_{4,1} + x_{4,3} + x_{4,4} + x_{4,6} + x_{5,1} + x_{5,2} + x_{5,4}$$

$$7925 = x_{1,2} + x_{2,1} + x_{2,3} + x_{2,6} + x_{3,1} + x_{3,3} + x_{3,5} + x_{3,7} + x_{4,1} + x_{4,3} + x_{4,4} + x_{4,6} + x_{5,1} + x_{5,2} + x_{5,4}$$

$$2813 = x_{1,4} + x_{2,5} + x_{2,6} + x_{2,7} + x_{3,5} + x_{3,6} + x_{3,7} + x_{3,8} + x_{4,4} + x_{4,5} + x_{4,6} + x_{4,7} + x_{5,2} + x_{5,3} + x_{5,4}$$

$$2247 = x_{1,3} + x_{2,2} + x_{2,3} + x_{2,7} + x_{3,1} + x_{3,4} + x_{3,6} + x_{3,7} + x_{3,8} + x_{4,2} + x_{4,3} + x_{4,4} + x_{4,7} + x_{5,1} + x_{5,3} + x_{5,4}$$

$$1650 = x_{2,4} + x_{3,2} + x_{3,3} + x_{3,4} + x_{3,8} + x_{4,1} + x_{4,2} + x_{4,3} + x_{4,5} + x_{4,6} + x_{4,7} + x_{5,1} + x_{5,2} + x_{5,1} + x_{5,3} + x_{5,4}$$

We now have 5 equations to solve linking the number of participants in the different $CHADS_2$ score groups (for ch_1 through to ch_5) to the possible combination subgroups ($x_{i,j}$) making up these groups along with the 5 equations relating the number with each $CHADS_2$ stroke risk factor to these same subgroups.

With 30 unknown numbers ($x_{1,1}$ to $x_{5,4}$) but only 10 equations it is not possible to solve these simultaneous equations analytically. Numerical optimisation can be used instead in which 'plausible' initial starting values for the $x_{i,j}$ are selected, with the values for the $x_{i,j}$ repeatedly adjusted until a solution is found. We are further limited by the numbers available in the CPRD Aurum apixaban trial-eligible cohort – that is there exists an upper bound to the values of the $x_{i,j}$ based on the number of participants in this subgroup in the data. For example, for $x_{4,2}$ (the number of patients with $CHADS_2$ score 4 having the risk factor combination CDS [CHF, diabetes, and prior Stroke or TIA]) there are only 27 patients in CPRD Aurum apixaban trial-eligible cohort with this combination of risk factors meaning our solution for $x_{4,2}$ is restricted to the integer range {0, 27}. Adding these restrictions aided discovery of potential solutions for the values of $x_{i,j}$ available within the CPRD

Aurum apixaban trial-eligible cohort, a potential solution for the $x_{i,j}$ values which was found via numerical optimisation is given in the Appendix.

Having found potential solutions considering only the $CHADS_2$ scores and combinations of risk factors, the next step was to consider additional important variables – namely sex, prior VKA exposure status, and a more refined breakdown of age. The distribution of stroke risk factors and age in the ARISTOTLE trial participants differed between men and women (133)with women older than men on average (median age 72 vs. 69) and a higher proportion of women having $CHADS_2$ score ≥3 (34.1% vs. 28.1% for men). This information was used to create separate equations for women and men.

A publication on ARISTOTLE results according to age gave the proportion of trial participants aged 75-80, 80-90, and 90+ (134). Given that age is an important predictor of stroke risk, bleeding risk, and mortality, it was decided to match the trial closely on age by splitting the age group into smaller categories. Combining all factors gave the following combinations to consider in creating subgroups:

- $CHADS_2$ score
- combination of stroke risk factors [C H A D S]
- age group [<65, ≥65 to <75, ≥75 to <80, ≥80 to <90, ≥90 years]
- sex
- prior VKA exposure

Multiplying the 32 possible combinations presented in Table A2.3.2.2.1 with these additional variables gave a total of 512 possible subgroups with a potential solution shown in Table A2.3.2.2.2.

SAS proc surveyselect was used to select a random sample from each subgroup with the size of the samples equalling the numbers found in the chosen solution. Specification of a different seed number within proc surveyselect would generate a different selection thus there are 2 potential sources of variation with this method: i) the solution used to satisfy the equations (many potential solutions to the simultaneous equations) and ii) the random sample selected for a given solution.

Table A2.3.2.2.2 Example solution for subgroup sampling

		Women		Men		Total	ARISTOTLE
Subgroup	**Stroke Risk Factors**	**CPRD Aurum (vka0/vka1)**	**example soln n (vka0/vka1)**	**CPRD Aurum (vka0/vka1)**	**example soln n (vka0/vka1)**	**soln n (vka0/vka1)**	**Apixaban (N = 9120)**
ch_0_age1	LVEF and/or SE	8/2	7/2	31/10	13/8	20/10	
ch_0_age2	LVEF and/or SE	8/5	7/3	23/19	1/13	8/16	
Total ch_0			**19 (14/5)**		**35 (14/21)**	**54 (28/26)**	**54**
ch_1_age1	C	40/24	22/19	172/77	81/67	189 (103/86)	
ch_1_age1	H	401/76	236/74	796/245	774/221	1305 (1010/295)	
ch_1_age1	D	74/14	1/7	169/32	1/4	13 (2/11)	
ch_1_age2	C	83/41	5/34	161/96	49/77	165 (54/111)	
ch_1_age2	H	1266/319	224/274	1664/487	350/467	1315 (574/741)	
ch_1_age2	D	80/21	0/1	202/53	0/2	3 (0/3)	
Total ch_1 young			897 (488/409)		2093 (1255/838)	2990 (1743/1247)	
ch_1_old0	A	400/130	10/16	423/147	3/9	38 (13/25)	
ch_1_old1	A	587/174	5/4	552/194	0/3	12 (5/7)	
ch_1_old2	A	92/20	0/0	73/23	0/0	0 (0/0)	
Total ch_1 old			15/20		15 (3/12)	50 (18/32)	
Total ch_1			**932 (503/429)**		**2108 (1258/850)**	**3040 (1761/1279)**	**3046**
ch_2_age1_RF3	CH	46/21	43/17	135/68	127/66	253 (170/83)	
ch_2_age1_RF2	CD	6/2	1/2	30/20	9/19	31 (10/21)	
ch_2_age1_RF4	HD	104/23	103/23	282/95	203/93	422 (306/116)	
ch_2_age1_RF1	S	66/19	6/16	151/49	48/45	115 (54/61)	
ch_2_age2_RF3	CH	143/71	138/63	214/175	161/156	518 (299/219)	
ch_2_age2_RF2	CD	18/12	3/3	37/39	3/28	37 (6/31)	
ch_2_age2_RF4	HD	337/109	139/102	653/220	207/210	658 (346/312)	
ch_2_age2_RF1	S	122/48	2/18	213/87	42/83	145 (44/101)	
Total ch_2 young			679 (435/244)		1500 (800/700)	2179 (1235/944)	
ch_2_old0	AC	61/47	18/24	74/60	7/43	92 (25/67)	
ch_2_old1	AC	134/90	2/10	120/125	0/11	23 (2/21)	
ch_2_old2	AC	34/30	0/0	36/20	0/1	1 (0/1)	
ch_2_old0	AH	872/277	102/168	845/308	46/238	554 (148/406)	
ch_2_old1	AH	1636/564	44/98	1034/417	37/174	353 (81/272)	

		Women		Men		Total	ARISTOTLE
Subgroup	**Stroke Risk Factors**	**CPRD Aurum (vka0/vka1)**	**example soln n (vka0/vka1)**	**CPRD Aurum (vka0/vka1)**	**example soln n (vka0/vka1)**	**soln n (vka0/vka1)**	**Apixaban (N = 9120)**
ch_2_old2	AH	315/120	1/1	133/56	1/2	5 (2/3)	
ch_2_old0	AD	61/22	4/6	84/44	2/7	19 (6/13)	
ch_2_old1	AD	81/33	1/3	84/55	0/1	5 (1/4)	
ch_2_old2	AD	11/4	0/0	18/7	0/0	0 (0/0)	
Total ch_2 old			482 (172/310)		570 (93/477)	1052 (265/787)	
Total ch_2			**1161 (607/554)**		**2070 (893/1177)**	**3231 (1500/1731)**	**3262**
ch_3_age1_RF1	CHD	7/6	2/5	11/12	10/12	29 (12/17)	
ch_3_age1_RF2	CS	9/6	0/1	23/6	3/5	9 (3/6)	
ch_3_age1_RF3	HS	58/16	48/14	127/47	22/46	130 (70/60)	
ch_3_age1_RF4	DS	22/16	18/14	62/53	21/51	104 (39/65)	
ch_3_age2_RF1	CHD	10/11	2/6	20/20	8/15	31 (10/21)	
ch_3_age2_RF2	CS	12/5	0/0	33/28	0/2	2 (0/2)	
ch_3_age2_RF3	HS	222/82	53/76	320/150	17/143	289 (70/219)	
ch_3_age2_RF4	DS	91/64	18/55	181/155	13/135	221 (31/190)	
Total ch_3 young			312 (141/171)		503 (94/409)	235/580	
ch_3_old0	ACH	128/113	51/71	139/114	21/93	236 (72/164)	
ch_3_old1	ACH	469/398	27/63	323/263	5/187	282 (32/250)	
ch_3_old2	ACH	179/94	0/0	83/68	1/3	4 (1/3)	
ch_3_old0	ACD	15/9	1/3	17/20	0/7	11 (1/10)	
ch_3_old1	ACD	30/28	2/2	25/34	1/2	7 (3/4)	
ch_3_old2	ACD	5/6	0/0	5/4	0/0	0 (0/0)	
ch_3_old0	AHD	241/134	18/40	376/124	8/95	161 (26/135)	
ch_3_old1	AHD	470/234	10/23	411/211	6/45	84 (16/68)	
ch_3_old2	AHD	69/41	0/0	55/22	0/1	1 (0/1)	
ch_3_old0	AS	74/38	6/16	119/63	2/33	57 (8/49)	
ch_3_old1	AS	150/85	2/5	176/96	2/13	22 (4/18)	
ch_3_old2	AS	37/12	0/0	32/17	0/0	0 (0/0)	
Total ch_3 old			340 (117/223)		525 (46/479)	865 (163/702)	
Total ch_3			**652 (258/394)**		**1028 (140/888)**	**1680 (398/1282)**	**1681**
ch_4_age1_RF1	SCD	1/3	0/1	5/3	0/2	3 (0/3)	

		Women		Men		Total	ARISTOTLE
Subgroup	**Stroke Risk Factors**	**CPRD Aurum (vka0/vka1)**	**example soln n (vka0/vka1)**	**CPRD Aurum (vka0/vka1)**	**example soln n (vka0/vka1)**	**soln n (vka0/vka1)**	**Apixaban (N = 9120)**
ch_4_age1_RF2	SCH	6/4	1/2	19/19	4/18	25 (5/20)	
ch_4_age1_RF3	SDH	19/8	1/6	48/27	1/26	34 (2/32)	
ch_4_age2_RF1	SCD	3/4	0/1	5/3	0/1	2 (0/2)	
ch_4_age2_RF2	SCH	35/30	0/7	44/52	3/41	51 (3/48)	
ch_4_age2_RF3	SDH	94/61	2/16	166/64	2/27	47 (4/43)	
Total ch_4 young			37 (4/33)		125 (10/115)	14/148	
ch_4_RF1_old0	ASC	11/8	2/6	15/20	1/9	18 (3/15)	
ch_4_RF1_old1	ASC	27/34	1/6	17/34	0/5	12 (1/11)	
ch_4_RF1_old2	ASC	9/11	0/0	7/13	0/0	0 (0/0)	
ch_4_RF2_old0	ASD	14/9	0/1	18/20	1/1	3 (1/2)	
ch_4_RF2_old1	ASD	23/18	2/2	29/26	0/2	6 (2/4)	
ch_4_RF2_old2	ASD	5/3	0/0	6/6	0/0	0 (0/0)	
ch_4_RF3_old0	ASH	224/122	33/56	249/139	8/109	206 (41/165)	
ch_4_RF3_old1	ASH	592/281	23/51	407/256	10/56	(140) 33/107	
ch_4_RF3_old2	ASH	144/110	4/3	75/36	0/1	8 (4/4)	
ch_4_RF4_old0	ACDH	103/82	15/26	108/105	4/72	117 (19/98)	
ch_4_RF4_old1	ACDH	230/219	15/25	194/179	2/51	93 (17/76)	
ch_4_RF4_old2	ACDH	42/50	1/0	30/31	0/1	2 (1/1)	
Total ch_4 old			272 (96/176)		333 (26/307)	605 (122/483)	
Total ch_4			**309 (100/209)**		**458 (36/422)**	**767 (136/631)**	**767**
ch_5_age1_RF	SCDH	7/5	1/3	15/18	1/10	15 (2/13)	
ch_5_age2_RF	SCDH	31/29	1/6	52/62	0/10	17 (1/16)	
Total ch_5 young			11 (2/9)		21 (1/20)	32 (3/29)	
ch_5_RF1_old0	ASCD	5/4	0/0	9/8	0/2	2 (0/2)	
ch_5_RF1_old1	ASCD	8/17	1/2	18/28	0/5	8 (1/7)	
ch_5_RF1_old2	ASCD	2/7	0/1	0/5	0/1	2 (0/2)	
ch_5_RF2_old0	ASCH	42/44	9/20	35/51	2/20	51 (11/40)	
ch_5_RF2_old1	ASCH	151/183	6/23	124/147	0/46	75 (6/69)	
ch_5_RF2_old2	ASCH	54/65	2/3	28/36	0/5	10 (2/8)	
ch_5_RF3_old0	ASDH	80/56	5/15	114/52	6/25	51 (11/40)	
ch_5_RF3_old1	ASDH	168/127	4/11	176/123	6/14	35 (10/25)	

		Women		Men		Total	ARISTOTLE
Subgroup	**Stroke Risk Factors**	**CPRD Aurum (vka0/vka1)**	**example soln n (vka0/vka1)**	**CPRD Aurum (vka0/vka1)**	**example soln n (vka0/vka1)**	**soln n (vka0/vka1)**	**Apixaban (N = 9120)**
ch_5_RF3_old2	ASDH	39/26	0/4	22/8	0/1	5 (0/5)	
Total ch_5 old			106 (27/79)		133 (14/119)	239 (41/198)	
Total ch_5			**117 (29/88)**		**154 (15/139)**	**271 (44/227)**	**273**
ch_6_RF0_old0	ASCDH	38/53	1/9	42/59	0/6	16 (1/15)	
ch_6_RF0_old1	ASCDH	70/103	1/8	75/105	0/9	18 (1/17)	
ch_6_RF0_old2	ASCDH	24/13	0/3	12/12	0/0	3 (0/3)	
Total ch_6			**22 (2/20)**		**15 (0/15)**	**37 (2/35)**	**37**
Total C							2784
Total H							7962
Total A			1257 (455/828)		1591 (156/1409)	2848 (611/2237)	**2850**
Total D							2284
Total S							1748
TOTAL			3212 (1513/1699)		5868 (2356/3512)	9080 (3869/5211)	**9120 (3912/5208)**

CPRD Aurum columns show the number of patients in the ARISTOTLE-eligible apixaban CPRD cohort available for selection.
example solution columns show example solutions of sample sizes that may be selected from the CPRD Aurum ARISTOTLE-eligible apixaban cohort to give a cohort matching the ARISTOTLE trial on key baseline characteristics.
Subgroup is defined by the CHADS2 score, combination of CHADS2 stroke risk factors, and age group with this further broken down by prior VKA exposure status within the table.
soln=solution; vka0 = VKA-naïve; vka1 = VKA-experienced; age1 = age < 65;

A2.3.3 Full procedure for the selection of the ARISTOTLE-analogous cohort

Step 1	Apply ARISTOTLE inclusion and exclusion criteria to the AF patients with exposure to apixaban and/or warfarin in CPRD Aurum. Since we will not select the index date for warfarin users at this stage we must apply the eligibility criteria at every single potential index date in the study period for each warfarin user.
Step 2	Select subset of apixaban users that match ARISTOTLE baseline characteristics using the random sampling within subgroups method. This step yields an ARISTOTLE-analogous apixaban arm.
Step 3	Order the ARISTOTLE-analogous apixaban arm in duration of prior exposure from shortest prior VKA exposure to longest prior VKA exposure.
Step 4	Match the VKA-naïve (the new users of apixaban) in the ARISTOTLE-analogous apixaban arm to the trial-eligible new users warfarin using propensity score matching.
Step 5	Remove any warfarin users selected as a match for the new users of apixaban form the pool of potential prevalent users.
Step 6	Order the pool of potential warfarin users (including all potential index dates) in duration of prior exposure from shortest prior VKA exposure to longest prior VKA exposure.
Step 7	Step through the prevalent apixaban users in the ARISTOTLE-analogous apixaban arm 1 by 1 in order of duration of prior VKA exposure. For each prevalent apixaban user select a sample of 5 warfarin users with equivalent duration of prior VKA exposure. Check the eligibility of the sampled warfarin users and if any are ineligible drop them from the sample and remove all other index dates belonging to the ineligible warfarin user from the pool of potential prevalent users. If any ineligible users were found, then sample additional prevalent warfarin users until 5 eligible are found.
Step 8	Categorise the prior VKA exposure into ‘treatment history strata’. Propensity score match the prevalent ARISTOTLE-analogous apixaban users 1:1 with prevalent warfarin users within the treatment history strata requiring that the match have equivalent prior VKA exposure.
Step 9	Set together the propensity score matched new and prevalent users giving the full ARISTOTLE-analogous CPRD Aurum cohort.

A2.3.4 Additional information of the analysis of outcomes

The outcomes used in the study all required hospitalisation or death and were captured using the HES and ONS datasets:

Effectiveness Outcomes

- Primary effectiveness outcome: stroke or systemic embolism
  including individual components (stroke, ischemic or uncertain type of stroke, haemorrhagic stroke, systemic embolism)
- Key secondary outcome: death from any cause
- Other secondary outcomes:
  Stroke, systemic embolism, or death from any cause
  Myocardial infarction
  Stroke, systemic embolism, myocardial infarction, or death from any cause
  Pulmonary embolism or deep-vein thrombosis

Safety Outcomes

- Primary safety outcome: Major bleeding
  including by location (intracranial, other location, gastrointestinal)
- Net clinical outcomes:
  Stroke, SE, or major bleeding
  Stroke, SE, major bleeding, or death from any cause

Cox proportional hazards models were used to analyse all time to event outcomes. The models were stratified by prior VKA exposure status (naïve/experienced). Cluster-robust standard errors were used with pair membership as the clustering variable (158, 159). The proportional hazards assumption was assessed by looking at the log-log of the Kaplan-Meier survival curves and inspection of scaled Schoenfeld residuals plotted against time. The primary analysis for the efficacy outcomes used an intent-to-treat censoring approach to mimic the primary analysis in the trial. Safety outcomes were analysed using an on-treatment censoring approach. Efficacy analyses were also analysed using an on-treatment censoring approach.

A2.3.5 Additional information on the prevalent new user design

Suissa in 2017 provided a framework for how to conduct pharmacoepidemiological studies including both new and prevalent users, a design he named the 'Prevalent New User' (PNU) design(137). The method Suissa proposed can be summarised as follows, supposing one is interested in comparing an old 'comparator' drug and a newer 'study drug':

1. **Identification of the 'base cohort'**, selection of all users of the comparator and newer study drug. This includes both patients newly initiating each treatment as well as patients switching from the comparator treatment to the newer study drug.
2. **Construction of exposure sets**. Suissa defines an exposure set as "the set of subjects in the base cohort exposed to the comparator drug at the point that a subject switched to the study drug". For each patient switching from the comparator to the study drug, the 'switchers', an exposure set is created comprising all patients continuing on the comparator treatment that have the same history of prior treatment.
3. **Calculation of time-conditional propensity scores**. All exposure sets are set together into 1 dataset. Conditional logistic regression is used to estimate the probability of switching to the study drug against the probability of continuing on the comparator within each exposure set.
4. **Selection of patients via propensity score matching in chronological order**. Having calculated the time-conditional propensity scores in step 3 we proceed to select patients into the final cohort in a process designed to mimic the selection into an RCT. In chronological order, from earliest index date of a study drug user to latest, select the closest user of comparator match to each study drug user. The positivity condition is checked by making sure the time-conditional propensity score of the switcher is within the range of time-conditional propensity scores of the patients in their exposure set. Should this condition not be met then the switcher is excluded

from the final cohort. Once a continuer has been selected as a match they are no longer eligible as a match at any other time and are therefore excluded from subsequent exposure sets when selecting a match.

After completion of these steps the resulting cohort should comprise a mix of new users of the study drug matched to new users of the comparator along with patients that switched from the compactor to the study drug matched to patients that continued on the comparator. The history of prior treatment may be defined based on a time metric such as number of days covered by the prior prescriptions or by the number of prior prescriptions. Depending on the typical use of the study drug and comparator drug of interest, use of a time metric that requires an exact match on the number of days prior exposure would be prohibitively restrictive and implausibly accurate for longer durations; selection of a suitable time interval such as ±1 month may therefore be more appropriate when assessing whether a comparator has ‘equivalent’ prior exposure to a switcher.

Appendix 3

A3.1 Supplementary material from Research Paper 2: PLOS Medicine results paper

S1 STROBE

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Title Abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction, paragraph 5; Methods of Analysis, Benchmarking results against ARISTOTLE
Methods			
Study design	4	Present key elements of study design early in the paper	Abstract, Methods and Findings, paragraph 1; Materials and methods, Study design
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Materials and methods, Setting/data sources
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up (*b*) For matched studies, give matching criteria and number of exposed and unexposed	Materials and methods, Patient Selection, Step 1 and Step 2; Table A2 in S3; Fig 2 Materials and methods, Patient Selection, Step 3; Fig 2; Results of Propensity score matching
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Materials and methods, Diagnostic and therapeutic codelists and Exposures and outcomes; Table 1;
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Exposures and outcomes; Table A2 in S3.
Bias	9	Describe any efforts to address potential sources of bias	Patient selection, Step 3; Methods of Analysis, Confounding and bias
Study size	10	Explain how the study size was arrived at	Fig 2;

			Protocol in S2.
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Introduction, paragraph 2 for TTR
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Patient selection, Step 3; Methods of Analysis, paragraph 1; Methods of Analysis, Sensitivity analyses
		(*b*) Describe any methods used to examine subgroups and interactions	Methods of analysis, Supplementary analyses
		(*c*) Explain how missing data were addressed	Methods of analysis, Missing data; Table 2 row 6 on prior INR
		(*d*) If applicable, explain how loss to follow-up was addressed	Methods of Analysis, paragraph 1; Methods of Analysis, Sensitivity analyses, paragraph 1.
		(*e*) Describe any sensitivity analyses	Methods, Sensitivity analyses.
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Fig 2
		(b) Give reasons for non-participation at each stage	Fig 2
		(c) Consider use of a flow diagram	Fig 2
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table 3, Table A8 in S3
		(b) Indicate number of participants with missing data for each variable of interest	Table 3
		(c) Summarise follow-up time (eg, average and total amount)	Main results, paragraph 1; Table A3 in S3, Table A5 in S3
Outcome data	15*	Report numbers of outcome events or summary measures over time	Tables A3 and A5 in S3

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Main results Table 3 shows categorisation of variables; Methods of analyses, Supplementary analyses Not appropriate for non-inferior results. Absolute event rates (%/yr) provided in Fig3 and Fig4.
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Results, Analysis of impact of warfarin time in therapeutic range (TTR); Results, Analysis of apixaban dose-adjustment; Results, Sensitivity analyses.
Discussion			
Key results	18	Summarise key results with reference to study objectives	Discussion paragraph 1;
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Sensitivity Analyses; Limitations
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Discussion paragraphs 1 and 2
Generalisability	21	Discuss the generalisability (external validity) of the study results	Limitations, paragraph 3; Conclusions paragraphs 1 and 2
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Funding statement

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

S2 ISAC Protocol

ISAC EVALUATION OF PROTOCOLS FOR RESEARCH INVOLVING CPRD DATA

FEEDBACK TO APPLICANTS

CONFIDENTIAL			***by e-mail***
PROTOCOL NO:	19_066R		
PROTOCOL TITLE:	Use of non-interventional data for determining the real-world effectiveness of anticoagulation medication for stroke prevention in a clinical trial analogous population		
APPLICANT:	Dr Kevin Wing London School of Hygiene & Tropical Medicine [redacted]		
APPROVED ☐	**APPROVED WITH COMMENTS (resubmission not required)** ☐	**REVISION/ RESUBMISSION REQUESTED** ☐	**REJECTED** ☐
INSTRUCTIONS: ***Protocols with an outcome of 'Approved' or 'Approved with comments' do not require resubmission to the ISAC.*** **REVIEWER COMMENTS:** **APPLICANT FEEDBACK:**			
DATE OF ISAC FEEDBACK:	19/09/19		
DATE OF APPLICANT FEEDBACK:			

For protocols approved from 01 April 2014 onwards, applicants are required to include the ISAC protocol in their journal submission with a statement in the manuscript indicating that it had been approved by the ISAC (with the reference number) and made available to the journal reviewers. If the protocol was subject to any amendments, the last amended version should be the one submitted.

Guidance on resubmitting applications, or making amendments to approved protocols, can be found on the CPRD website at https://cprd.com/research-applications.

Medicines & Healthcare products Regulatory Agency

CPRD

INDEPENDENT SCIENTIFIC ADVISORY COMMITTEE (ISAC) PROTOCOL APPLICATION FORM

PART 1: APPLICATION FORM

IMPORTANT

Both parts of this application must be completed in accordance with the guidance note 'Completion of the ISAC Protocol Application Form', which can be found on the CPRD website cprd.com/research-applications

FOR ISAC USE ONLY	
Protocol No. -	**Submission date -**

GENERAL INFORMATION ABOUT THE PROPOSED RESEARCH STUDY

1. Study Title (Max. 255 characters)

Use of non-interventional data for determining the real-world effectiveness of anticoagulation medication for stroke prevention in a clinical trial analogous population

2. Research Area (place '**X**' in all boxes that apply)

Drug Safety	X	Economics	
Drug Utilisation		Pharmacoeconomics	
Drug Effectiveness	X	Pharmacoepidemiology	X
Disease Epidemiology		Methodological	X
Health Services Delivery			

3. Chief Investigator

Title:	Dr
Full name:	Kevin Wing
Job title:	Assistant Professor of Epidemiology
Affiliation/organisation:	London School of Hygiene & Tropical Medicine
Email address:	[redacted]
CV Number (if applicable):	

4. Corresponding Applicant

Title:	Ms
Full name:	Emma Powell
Job title:	Research Degree Student
Affiliation/organisation:	London School of Hygiene & Tropical Medicine
Email address:	[redacted]
CV Number (if applicable):	

5. List of all investigators/collaborators

Title:	Ms
Full name:	Emma Powell
Job title:	Research Degree Student
Affiliation/organisation:	London School of Hygiene & Tropical Medicine
Email address:	[redacted]
CV Number (if applicable):	
Will this person be analysing the data? (Y/N)	Y

Title:	Dr
Full name:	Kevin Wing
Job title:	Assistant Professor of Epidemiology
Affiliation/organisation:	London School of Hygiene & Tropical Medicine
Email address:	[redacted]
CV Number (if applicable):	
Will this person be analysing the data? (Y/N)	N

Title:	Dr
Full name:	Ian Douglas
Job title:	Associate Professor of Pharmacoepidemiology
Affiliation/organisation:	London School of Hygiene & Tropical Medicine
Email address:	[redacted]
CV Number (if applicable):	
Will this person be analysing the data? (Y/N)	N

Title:	Ms
Full name:	Usha Gungabissoon
Job title:	Director, Epidemiology
Affiliation/organisation:	GlaxoSmithKline
Email address:	[redacted]
CV Number (if applicable):	
Will this person be analysing the data? (Y/N)	N

Title:	Prof
Full name:	Liam Smeeth
Job title:	Professor of Clinical Epidemiology
Affiliation/organisation:	London School of Hygiene & Tropical Medicine
Email address:	[redacted]
CV Number (if applicable):	
Will this person be analysing the data? (Y/N)	N

[Add more investigators/collaborators as necessary by copy and pasting a new table for each investigator/collaborator]

6. Experience/expertise available

List below the member(s) of the research team who have experience with CPRD data.

Name:	**Protocol Number/s:**
Kevin Wing, Ian Douglas, Usha Gungabissoon, Liam Smeeth	>50 protocols

List below the member(s) of the research team who have statistical expertise.

Name(s):
Emma Powell

List below the member(s) of the research team who have experience of handling large datasets (greater than 1 million records).

Name(s):
Kevin Wing
Ian Douglas

List below the member(s) of the research team, or supporting the research team, who have experience of practicing in UK primary care.

Name(s):
Liam Smeeth

ACCESS TO THE DATA

7. Sponsor of the study

Institution/Organisation:	Medical Research Council
Address:	Polaris House, North Star Avenue, Swindon, SN2 1FL, United Kingdom

8. Funding source for the study

Same as Sponsor?	Yes	X	No	
Institution/Organisation:	Medical Research Council			
Address:	Polaris House, North Star Avenue, Swindon, SN2 1FL, United Kingdom			

9. Institution conducting the research

Same as Sponsor?	Yes		No	X
Institution/Organisation:	London School of Hygiene & Tropical Medicine			
Address:	Keppel Street, London, WC1E 7HT, United Kingdom			

10. Data Access Arrangements

Indicate with an '**X**' the method that will be used to access the data for this study:

Study-specific Dataset Agreement	

Institutional Multi-study Licence	X
Institution Name	London School of Hygiene & Tropical Medicine
Institution Address	Keppel Street, London, WC1E 7HT, United Kingdom

Will the dataset be extracted by CPRD?

Yes		No	X

If yes, provide the reference number:

11. Data Processor(s):

Processing	X

Accessing	X	
Storing	X	
Processing area (UK/EEA/Worldwide)		UK
Organisation name		London School of Hygiene & Tropical Medicine
Organisation address		Keppel Street, London, WC1E 7HT, United Kingdom

Processing		
Accessing		
Storing		
Processing area (UK/EEA/Worldwide)		
Organisation name		
Organisation address		

[Add more processors as necessary by copy and pasting a new table for each processor]

INFORMATION ON DATA

12. Primary care data (place '**X**' in all boxes that apply)

CPRD GOLD	X	CPRD Aurum	X

13. Please select any linked data or data products being requested

Patient Level Data (place '**X**' in all boxes that apply)

ONS Death Registration Data	X	CPRD Mother Baby Link	
HES Admitted Patient Care	X	Pregnancy Register	
HES Outpatient		NCRAS (National Cancer Registration and Analysis Service) Cancer Registration Data	
HES Accident and Emergency		NCRAS Cancer Patient Experience Survey (CPES) data	
HES Diagnostic Imaging Dataset		NCRAS Systemic Anti-Cancer Treatment (SACT) data	
HES PROMS (Patient Reported Outcomes Measure)		NCRAS National Radiotherapy Dataset (RTDS) data	
		Mental Health Services Data Set (MHDS)	

Area Level Data (place '**X**' in all boxes that apply)

Practice level (UK)		**Patient level (England only)**	
Practice Level Index of Multiple Deprivation (Standard)	X	Patient Level Index of Multiple Deprivation	X
Practice Level Index of Multiple Deprivation (Non-standard)		Patient Level Townsend Score	
Practice Level Index of Multiple Deprivation Domains (Non-standard)			
Practice Level Carstairs Index for 2011 Census (Excluding Northern Ireland) (Standard)			
2011 Rural-Urban Classification at LSOA level (Non-standard)			

Reference number (where applicable):

14. Are you requesting linkage to a dataset not listed above?

Yes		No	X

If yes, provide the reference number:

15. Does any person named in this application already have access to any of these data in a patient identifiable form, or associated with an identifiable patient index?

Yes		No	X

If yes, provide further details:

VALIDATION/VERIFICATION

16. Does this protocol describe an observational study using purely CPRD data?

Yes	X	No	

17. Does this protocol involve requesting any additional information from GPs, or contact with patients?

Yes		No	X

If yes, provide the reference number:

Medicines & Healthcare products Regulatory Agency

CPRD

PART 2: PROTOCOL INFORMATION

Applicants must complete all sections listed below **Sections which do not apply should be completed as '*Not Applicable*' and justification provided**
A. Study Title (Max. 255 characters) Use of non-interventional data for determining the effectiveness of anticoagulation medication for stroke prevention
B. Lay Summary (Max. 250 words) Atrial Fibrillation (AF) is a heart condition in which patients have an irregular heartbeat. Patients with AF are at a higher risk of stroke and may be prescribed a type of medication called anticoagulants to reduce the risk of stroke. Treatment guidelines for anticoagulants are based on the results of randomised clinical trials which have very strict entry criteria. This means that many people with AF who are prescribed these drugs by their GP could be quite different to patients studied in clinical trials. The types of patients who are usually not allowed to take part in clinical trials include people with existing medical conditions, people with no other stroke risk factors, and pregnant patients. Patients included in anticoagulant clinical trials may therefore not be representative of the patients who are prescribed these drugs in clinical practice. In this study we will look at how well a specific anticoagulant medicine called apixaban works compared with another treatment called warfarin in people who would have been excluded from a landmark anticoagulant trial (the ARISTOTLE trial), in order to help improve treatment guidelines for these people. We will use information routinely collected by GPs and hospitals to assess the impact of apixaban and warfarin in preventing stroke and blood vessel blockage. First we will see how well these treatments work when prescribed to people that are similar to those included in the ARISTOTLE trial. Then we will see how well the treatments work when prescribed to patient groups excluded from this trial.
C. Technical Summary (Max. 300 words) Patients with atrial fibrillation (AF) are at a greatly increased risk of stroke; prophylactic treatment with anticoagulation medication reduces this risk. In the last decade several direct oral anticoagulants (DOACs) have been approved providing an alternative to the standard treatment warfarin which has many drug and dietary interactions and requires onerous monitoring. Treatment guidelines for AF patients are based on the results from randomised controlled trials (RCT). There is increasing interest in the effectiveness of medications in routine clinical practice to confirm trial results and estimate drug effectiveness in patient groups excluded from or underrepresented in clinical trials. There is however some uncertainty about the suitability of using non-interventional data to address questions about drug effectiveness and on the most suitable methods to be used. The aims of this study are to attempt to measure the association between anticoagulation treatments for stroke prevention in AF using electronic health records (EHRs) and to develop a methodological framework for using observational EHRs to answer questions about DOACs among patients excluded from or underrepresented in the RCTs. This study will use individual patient data from ARISTOTLE[1], a pivotal trial conducted 2006-2011 in 18,201 patients that demonstrated superiority of the DOAC apixaban compared with warfarin in prevention of stroke. The individual patient data will be used to match to UK NHS patients with anonymised routinely-collected EHRs from CPRD. Analysis of drug effectiveness in this cohort will help determine whether EHR data are suitable for this kind of research question. Selecting EHR patients similar to trial patients will remove much of the variability in baseline risk of the study outcomes between trial participants and EHR patients. If we can demonstrate replication of trial results in our ARISTOTLE-analogous cohort, the same methodology will be used to determine drug effectiveness in patient groups underrepresented in the trial.
D. Outcomes to be Measured Time to event: stroke, systemic embolism, myocardial infarction, all-cause death, major bleeding

E. Objectives, Specific Aims and Rationale

Aim 1: To measure the association between anticoagulation treatments for stroke prevention in AF and time to stroke, systemic embolism, myocardial infarction, major bleeding, and mortality amongst an ARISTOTLE-analogous cohort of patients from UK electronic health records (EHR).

Aim 2: To develop a methodological framework with in-built validation, for using observational electronic health records to answer questions about DOAC risks and benefits in patients excluded from or underrepresented in the RCTs.

Objectives

Objective 1. Check comparability of EHR data and robustness of methods for measuring AF stroke prevention medication effectiveness in EHR data by comparing with ARISTOTLE results.

We will obtain fully anonymised individual patient data from the pivotal ARISTOTLE trial. A group of individuals with similar characteristics will be selected from EHR data based on medical history, prescription data, and baseline characteristics. The ARISTOTLE trial measured the efficacy of apixaban vs warfarin for the prevention of stroke, systemic embolism, myocardial infarction, and mortality and safety (major bleeding) amongst people with AF and at least one risk factor for stroke. Individual trial participants will be matched with similar people in the anonymised EHR databases the Clinical Practice Research Datalink (CPRD) Gold and CPRD Aurum in order to create an ARISTOTLE-analogous cohort within EHRs. The individual trial participant information is being obtained from Bristol Myers-Squibb. Using cohort methodology, estimates of the effect of apixaban vs warfarin on stroke, systemic embolism, myocardial infarction, major bleeding and mortality in the CPRD cohort of ARISTOTLE analogous patients will then be measured. The results will be compared with the ARISTOTLE findings to determine the utility of CPRD records for measuring medication effectiveness in AF. This objective will provide a methodological framework for measuring drug effectiveness in people with AF, using observational data from CPRD.

Objective 2. Extension of trial findings: Measure AF treatment effects in patients excluded from ARISTOTLE

Using the methodological template developed in Objective 1, we will determine the effect of apixaban vs warfarin in prevention of stroke, systemic embolism, myocardial infarction, major bleeding and mortality in people with AF not eligible for the ARISTOTLE study (most importantly people with substantial comorbidity), and look separately at important subgroups e.g. those with and without underlying cardiovascular disease.

Objective 3. Comparative effectiveness: Compare treatment effectiveness between multiple individual anticoagulants in all anticoagulant recipients (no eligibility criteria other than diagnosis of AF)

Using the methodological template developed in Objective 1, we will compare time to stroke, systemic embolism, myocardial infarction, major bleeding, and death based on prescribed treatments:
a) warfarin
b) apixaban
c) rivaroxaban
d) dabigatran
Apixaban will be compared with warfarin then all other DOACs compared with apixaban.

Rationale

AF treatment guidelines are largely informed by randomised controlled trial (RCT) results, but we do not know if these findings apply to large patient populations not studied in trials. Apixaban is one of the most widely used anticoagulants used in stroke prevention in AF. It was studied in a large randomised trial (ARISTOTLE), but we don't know the effects of treatment in important patient groups who were not studied. Some were excluded from AF stroke prevention trials in general (e.g. those with mechanical heart valves and those with substantial comorbidity) and some are under-represented (e.g. elderly patients), meaning conclusions about these groups are difficult to make. The results we generate will firstly tell us if EHR data are suitable for this purpose. If so, our findings will aid patients, prescribers and policy makers in better understanding the benefits and risks of different anticoagulants for stroke prophylaxis in AF rather than assuming that the trial estimates are applicable to all patients.

F. Study Background

Atrial fibrillation (AF) is a heart condition in which the patient experiences a rapid and irregular heartbeat caused by electrical signals firing from multiple locations in the atria. Although patients may be asymptomatic , symptoms such as heart palpitations, fainting, light-headedness, and shortness of breath are reported. The prevalence of AF is estimated to be around 3%[2] and increases with age from 0.2% in people aged 45-54 years to 8.0% in those 75 and older[3]. The lack of organised atrial contraction in AF can lead to blood stagnating in the left atrium or left atrial appendage and the formation of thrombi. Should a thrombus move from the heart through the body this can cause systemic embolism or stroke; consequently patients with AF have a five fold higher risk of stroke. It is recommended that AF patients receive prophylactic treatment with anticoagulation medication to reduce the risk of stroke. The previous standard anticoagulation treatment for this indication, warfarin, has many treatment interactions and requires frequent monitoring and dose adjustments to stay within the therapeutic range of anticoagulant action as summarised by Hirsh et al[4]: Warfarin has interactions with a wide range of drugs such as metronidazole which inhibits warfarin clearance, barbiturates and carbamazepine which increase hepatic warfarin, and aspirin which increases the risk of bleeding. Diet also interacts with response to warfarin with increased intake of vitamin K (present in green vegetables) leading to a reduction in the anticoagulant response to warfarin. Genetics influences the warfarin dose-response relationship most notably in common mutations in coding for cytochrome P450 (the family of enzymes responsible for warfarin metabolism). Warfarin therapy is monitored by calculating a patient's International Normalised Ratio (INR), a standardised measurement of the time taken for blood to clot. Typically for AF the patient must maintain INR at a therapeutic range between 2.0 and 3.0 with INR values below 2 putting a patient at higher risk of stroke and levels above 3 resulting in a higher risk of bleeding. On initiation of therapy INR is checked daily until in therapeutic range, then 3 times weekly for 2 weeks, then less often, according to the stability of the results. Given the challenge in maintaining INR in therapeutic range and the complex safety profile of warfarin it was hoped that the introduction of the direct acting oral anticoagulants (DOACs) would provide a safer and easier to manage long term anticoagulation therapy for AF patients. The pivotal trial of the DOAC apixaban for this indication, ARISTOTLE, demonstrated superiority over warfarin for both the primary efficacy (prevention of stroke) and safety (major bleeding) outcomes.

Apixaban was licensed based on the results of the pivotal trial ARISTOTLE, a randomised controlled trial (RCT). ARISTOTLE had eligibility criteria that patients had to meet to be included in the trial, thus limiting the generalisability of the results of the trial. As a result, evidence on treatment effect is lacking for patients who would not have met the ARISTOTLE eligibility criteria such as individuals with a mechanical heart valve, those at increased bleeding risk, and individuals with severe comorbid conditions. The regulatory environment now demands evidence of treatment effectiveness outside the confines of randomised trials. Non-interventional data sources have the potential to overcome many of the RCT limitations given that they contain data for a wide spectrum of patients treated with the drug in routine care including patients who would have been excluded from trials. Data collected as a standard part of patient care such as electronic healthcare record (EHRs) provide a valuable opportunity to obtain evidence on the effectiveness of apixaban in a routine care setting. A key problem with non-interventional studies using these data is that the absence of randomisation leaves them highly susceptible to confounding (with confounding by indication a particular problem), making it difficult to have confidence in the results. By contrast matching to individual patient data from ARISTOTLE and then using novel methods for matching within EHR treatment groups should result in an EHR population similar to the trial population that is well balanced by treatment group. If successful, the estimates of effectiveness and safety of apixaban obtained from this approach should then be comparable with the ARISTOTLE results. If non-interventional data can be successfully used to approximate the findings of ARISTOTLE then they may be reliable to estimate effects in under studied AF patient groups. This project will involve testing whether EHR data can find results compatible with the ARISTOTLE trial results while developing optimal methodology for studying anticoagulants in stroke prophylaxis. This methodology can then be applied to under studied AF patient groups.

G. Study Type

Hypothesis testing

H. Study Design

This is a historical cohort study.

The cohort study design allows measurement of the effects of prescribing apixaban vs warfarin for prevention of stroke and systemic embolism in AF on key efficacy and safety outcomes..

I. Feasibility counts ARISTOTLE inclusion and exclusion criteria were extracted from the trial protocol. Read code and medication codelists were created for the inclusion and exclusion criteria. The criteria were applied to a January 2018 extract of CPRD Gold patients prescribed apixaban. Overall out of 13 332 patients with a prescription for apixaban and diagnosis of AF 63% (8 407) were trial-eligible. Trial criteria were also applied to patients with a prescription for warfarin in the period 01 January 2013 to 31 January 2018 and a diagnosis of AF (68 113 patients). Of these patients 45 435 (62.3%) were eligible according to trial criteria.
J. Sample size considerations We will include all eligible patients registered in the CPRD and who meet the trial criteria. In ARISTOTLE there were 9120 subjects in the apixaban arm therefore it was estimated a minimum of 15,000 EHR patients exposed to apixaban were needed for matching to be feasible. It was unlikely there would be enough patients in CPRD Gold alone for the project given that only ~8400 patients were eligible in the January 2018 extract. The CPRD Aurum database (June 2019 extract) contained 29,578 patients with both an atrial fibrillation diagnosis and a prescription for apixaban; of these patients 23,526 were not registered in practices that had previously contributed data to CPRD Gold. Using the assumption that the proportion of Aurum patients who would meet the ARISTOTLE trial eligibility criteria would be similar to the proportion of Gold patients (~60%) gave an estimate of 14,115 trial eligible Aurum patients in the apixaban arm. Combining the Gold and Aurum cohorts is therefore estimated to give >22,000 unique trial-eligible EHR apixaban patients.
K. Planned use of linked data (if applicable): We intend to use CPRD data linked with HES in patient data to enable optimal stroke, MI, and major bleeding ascertainment, and ONS mortality data to determine deaths. When matching CPRD patients we plan to include the practice level deprivation level index as a matching variable because socioeconomic status is predicted to influence the likelihood of the primary study outcome of stroke.

L. Definition of the Study population

For all objectives two datasets of UK primary care data will be combined: CPRD Gold and CPRD Aurum. Patients with a prescription for an oral anticoagulant in the time period 01 January 2013 to 31 January 2019 (exact cut-off date dependent on date of final data extracts) and a prior diagnosis of atrial fibrillation will be selected as the EHR cohort. Apixaban gained UK marketing authorisation in January 2013 for the indication of prevention of stroke and systemic embolism in AF patients; the minimum date is set to capture all patients prescribed apixaban for this indication in the UK. A previous study validated the use of diagnostic read codes for identifying patients with AF in CPRD by sending surveys to GPs and found a confirmation rate of 98% among patients originally identified with AF codes[5].

Objective 1

Step 1: We will select all (HES and ONS linked) patients in the EHR cohort who would have met the following eligibility criteria for inclusion in the ARISTOTLE study, at least 6 months after patient registration in an up to standard practice: or 6 months post-UTS date, whichever is later:

- a diagnosis of AF,
- age over 18 years,
- at least one of the following risk factors for stroke: diagnosis of congestive heart failure, hypertension requiring pharmacological treatment, age greater than 75, diabetes mellitus, and prior stroke or systemic embolism,
- no AF due to reversible causes (e.g. thyrotoxicosis, pericarditis),
- no clinically significant (moderate or severe) mitral stenosis, in EHRs clinical significance and severity are not consistently recorded therefore a diagnosis of mitral stenosis will be sufficient to exclude a patient,
- no increased bleeding risk that is believed to be a contraindication to oral anticoagulation (e.g. previous intracranial haemorrhage),
- no conditions other than AF that require chronic anticoagulation (e.g. prosthetic mechanical heart valve),
- no persistent, uncontrolled hypertension (systolic BP > 180 mm Hg, or diastolic BP > 100 mm Hg), in EHRs this will be implemented by excluding patients whose latest blood pressure reading in the 6 months prior to the index date is over the systolic or diastolic blood pressure limit
- no active infective endocarditis,
- no concomitant treatment with aspirin > 165 mg/day,
- no simultaneous treatment with both aspirin and a thienopyridine (e.g., clopidogrel, ticlopidine),
- no severe comorbid condition with life expectancy of ≤ 1 year,
- no active alcohol or drug abuse, or significant psychosocial difficulties (e.g., psychosis, dementia),
- no recent ischemic stroke (within 7 days),
- no severe renal insufficiency (serum creatinine > 2.5 mg/dL or a calculated creatinine clearance < 25 mL/min),
- no ALT or AST > 2X ULN or a Total Bilirubin ≥ 1.5X ULN (unless an alternative causative factor such as Gilbert's syndrome is identified),
- no platelet count ≤ 100,000/ mm^3,
- no haemoglobin < 9 g/dL,
- no prior exposure to apixaban,
- no women who are pregnant or breastfeeding

For the 3 criteria involving patient laboratory results (renal insufficiency, low platelet count, and low haemoglobin) a patient will be excluded if their last test result in the 90 days prior to the index date meets the exclusion criteria.

Step 2: Next we will determine if/when these patients received apixaban or warfarin. Individuals in EHR who have more than one warfarin eligibility period within their record will be able to contribute more than once to the pool of warfarin subjects (with the covariates and person-time contributed unique to the specific eligibility period) as long as they have no past apixaban exposure.

Step 3: Having obtained individual level patient data for ARISTOTLE participants from Bristol Myers-Squibb we will then match each ARISTOTLE apixaban participant 1:1 with the closest available apixaban patient record in our EHR pool. We will consider matching on a selection of the following ARISTOTLE baseline characteristics:

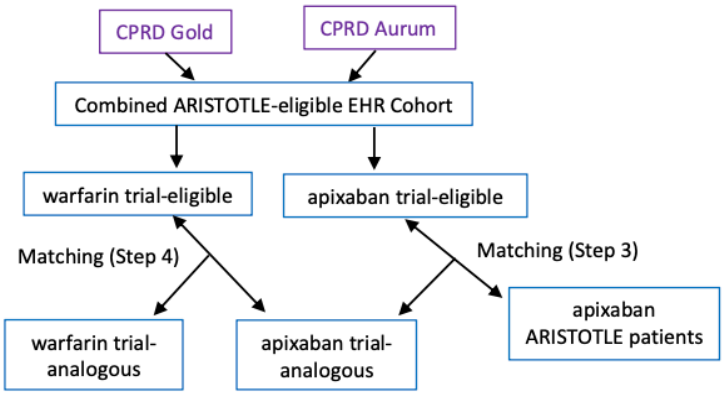
- age
- sex
- body mass index
- systolic blood pressure (mmHg)
- history of congestive heart failure or left ventricular systolic dysfunction

- hypertension requiring pharmacological treatment
- diabetes mellitus
- prior stroke or TIA or thromboembolism
- smoking status [current smoker, former smoker, or never smoked]
- alcohol consumption [none, low, moderate, heavy]
- renal impairment [severe, moderate, mild, normal based on the CrCL value]
- prior VKA/warfarin use
- concomitant treatment with:
  - anticoagulant/VKA use other than warfarin or apixaban
  - aspirin
  - antiplatelets
  - NSAIDs
  - lipid lowering drug therapy
  - CYP3A4 inhibitors

We anticipate matching all or the majority of ARISTOTLE apixaban subjects with an EHR patient, giving us a pool of ARISTOTLE-analogous apixaban patients, with similar baseline characteristics as ARISTOTLE subjects at the point of randomisation (n~9,000).

The variables selected as potential matching variables are those known or suspected to influence the likelihood of the outcomes of interest. The exact selection of matching variables will depend on the quality and completeness of the data available and a balance will be struck between the matched sample size and sample balance. Variables may be grouped together to increase the sample size, for example by grouping concomitant medications that increase bleeding risk. Continuous variables may be coarsened by splitting into appropriate categories. A procedure will be employed to facilitate selection of a matched cohort, for example via coarsened exact matching[6] for example by use of the %CEM SAS macro[7]. Coarsened Exact Matching is a nonparametric matching method that has been found to give estimates of casual effects with lower variance and bias for a given sample size compared with other commonly used methods of matching[8].

Figure 1: Assembly of Matched Trial-analogous Cohort of EHR Patients



Step 4: The resulting trial matched sample of EHR apixaban exposed subjects will be matched to the warfarin ARISTOTLE-eligible EHR subjects (Figure 1) using a matching method such as propensity score matching (PSM), or CEM (with the final method selected based upon method giving the optimal sample size versus balance). Where an individual from EHR has multiple warfarin "eligibility periods" that can be matched to an apixaban trial matched subject, the EHR characteristics that will be matched on will be those from the beginning of the specific eligibility period. The covariates for consideration in the matching between EHR treatment arms or construction of a PS model will include the variables listed above used in in step 3 along with additional EHR variables such as data source (CPRD Gold or Aurum), socioeconomic status, and comorbidities. The hazard ratio for the outcomes of interest (listed in section N) will then be calculated. Each apixaban patient from the ARISTOTLE-eligible EHR

patients will be matched 1:1 with the warfarin EHR patient with the closest match giving a trial-analogous analysis cohort of ~18,000.

A patient may be exposed to warfarin followed by apixaban in the time period of interest and the patient be trial eligible in both treatment periods; in such a situation both patient treatment periods may be included in the EHR treatment groups with the restriction that a patient must not be matched to themselves. Given the trial exclusion criteria that a patient may not have been previously exposed to apixaban, any warfarin treatment periods after apixaban exposure would not be eligible for inclusion in the warfarin EHR cohort.

Objective 2: we will select patients who would not have been included in the ARISTOTLE trial (and therefore would not have been included in the Objective 1 cohort) based on their age, stroke risk factors, or presence of substantial comorbidity. Specifically, this will be patients with an AF diagnosis in the EHR cohort meeting these additional criteria:

- age >77 years (although elderly patients were not excluded the maximum age at first dose in ARISTOTLE was 76),

OR

- no evidence of at least one additional risk factor for stroke

OR

- AF due to reversible causes

OR

- evidence of drug/alcohol abuse

OR

- severe comorbid condition: ARISTOTLE required patients to be excluded from the trial if they had serious disease with a likelihood of causing death within 1 year or reasons making participation unpractical (such as dementia).

In these special patient populations the same outcomes as objective 1 will be assessed as described in section N.

Objective 3: we will select all patients with AF with a prescription for any anticoagulant in the treatment period from the set of treatments: apixaban, warfarin, rivaroxaban, dabigatran. For this objective all outcomes listed in section N will be assessed. Patients will be stratified on whether they would have met the ARISTOTLE trial criteria.

M. Selection of comparison group(s) or controls

For objectives 1 and 2 the comparison group consists of the patients prescribed warfarin with a diagnosis of AF meeting the eligibility conditions described in section L. For objective 3 apixaban is compared to each other treatment group (warfarin, rivaroxaban, dabigatran).

N. Exposures, Outcomes and Covariates

The exposures of interest are apixaban, warfarin, rivaroxaban, and dabigatran (any dose for each exposure).
The individual effectiveness outcomes:

- stroke
- systemic embolism
- myocardial infarction
- all cause death

The safety outcome for the study:

- major bleeding (bleeding requiring transfusion, bleeding at a critical site, bleeding requiring attendance at hospital, or fatal bleeding)

For all outcomes we make the assumption that if there is no record of an outcome then the outcome did not occur.

The EHR ARISTOTLE-eligible patients prescribed apixaban will be matched to the ARISTOTLE apixaban patients on the variables listed in section L (step 3 of objective 1), with value taken at baseline (value at time of first dose or latest measurement/data recorded prior to first dose).

When matching between the treatment arms within the EHR cohort of patients the variables to be considered for inclusion in the matching algorithm include those listed in section L used in matching from the EHR treatment arm to the trial patient data in addition to EHR variables such as data source (CPRD Gold or Aurum), socioeconomic status, and comorbidities.

O. Data/ Statistical Analysis

Data mapping

For all 3 objectives the data source will consist of two datasets of UK real world clinical data which will be combined: CPRD Gold and CPRD Aurum. The two datasets will be mapped to a common data model based on adapted Clinical Data Interchange Standards Consortium (CDSIC) standards[9]. Duplicate entries of patients – where a patient is present in more than one dataset – will be removed with the following rule: keep in CPRD Aurum above CPRD Gold. CPRD Aurum provides a dataset listing practices in CPRD Aurum which have previously contributed data to CPRD Gold; this dataset will be used to exclude data from practices with eligible patients in CPRD Gold where the patient data is also recorded in Aurum. Only data of interest to the study will be mapped and used: patient demographics, diagnoses, clinical events, and recorded symptoms, therapies prescribed, lab results, and vital signs.

Primary Analysis

o Population

The ARISTOTLE trial used an intent to treat (ITT) approach for the primary efficacy analysis, and an on-treatment approach for the sensitivity analyses and safety outcomes. To perform an equivalent analysis with the EHR data the following analysis populations will be used for all objectives:

Prescribed Population: all patients who were prescribed a treatment, regardless of future changes to treatment. When summarizing data using this population, subjects are categorized according to the As Prescribed group.

On-treatment Population: all patients who were prescribed a treatment. In the case of patients discontinuing or switching treatment, data will be included up to and including their derived date of last dose of the initially prescribed treatment.

o Censoring

Index date

The index date for the EHR cohort will be the date of the patient's first prescription of apixaban or warfarin (for objectives 1 and 2), or for objective 3 first prescription of apixaban, warfarin, rivaroxaban, or dabigatran on or after the date the patient first met the eligibility criteria for the trial in the treatment period of interest

Date of last dose (all objectives)

The date of last dose will be estimated using the subject's date of prescription, number of tablets prescribed, and daily dose. Where there are missing values for the number of tablets prescribed or daily dose suitable values will be used to replace these, for example by substituting with the median or modal value. To allow for stockpiling of tablets and less than 100% adherence we will add 30 days after the apparent end of treatment when deriving date of last dose.

Primary censoring scheme:

Patients will be censored at the earliest of: outcome of interest, death date, 'transferred out date', 'last collection date', or 5 years after the index date. The 5-year limit reflects the maximum possible follow-up for a subject in the ARISTOTLE trial. Conclusions regarding noninferiority or superiority will be based on the results of the analyses using this censoring scheme as this most closely resembles the ARISTOTLE analysis plan.

Supportive censoring scheme:

Patients will be censored at the earliest of: outcome of interest, death date, 'transferred out date', 'last collection date', 5 years after the index date, or the derived last date of study drug. By censoring around the time of last study drug this scheme should include only events likely to be due to the drug taken.

o Primary outcome

The primary efficacy endpoint will be the time to first occurrence of confirmed stroke (ischemic, hemorrhagic, or of unspecified type), or systemic embolism during the study, regardless of whether the subject is receiving treatment at the time of the event (i.e. using the primary censoring scheme).

Comparisons will be made according to prescribed treatment (apixaban vs warfarin) for time to stroke/SE.

This analysis approach will be used as the primary analysis for all 3 objectives.

o Descriptive analyses

Demographic and Baseline Characteristics

Frequency distribution and summary statistics for demographic and baseline variables will be presented by treatment group for the EHR patients, both before and after matching steps.

Treatment Switching and Discontinuation

As the primary analysis accounts neither for treatment switching nor for treatment discontinuation and does not capture that these may be unfavourable outcomes both the proportion of patients discontinuing treatment and time to treatment discontinuation will be tabulated by prescribed treatment.

o Regression model
Hazard ratio (apixaban/warfarin) comparing the event rate (%/yr).
Interpretation: This estimand targets the treatment-policy effect of treatment initiation of apixaban vs initiation or continuation of warfarin on the time to first stroke or systemic embolism.
All time to event endpoints will be analysed using a Cox proportional hazards model including treatment group as a covariate and prior warfarin/VKA status (experienced, naïve).
Point estimates and two-sided 95% Cis for HR will be constructed for the outcome.

o Validation of Observational Results Against Aristotle Data
In Objective 1 alone we will validate the findings from our primary analysis against ARISTOTLE by determining whether results of the EHR analysis are compatible with the ARISTOTLE trial results. The ARISTOTLE trial demonstrated superiority of apixaban over warfarin for the primary endpoint (HR 0.79, 95% CI 0.66-0.95). The treatment effect seen with the EHR data may be weaker than that seen in ARISTOTLE.

A subgroup analysis looking at the outcomes of the EU patients in ARISTOTLE showed a smaller treatment difference with the estimate for HR below 1 but upper limit of the CI crossing 1 for the primary efficacy endpoint and death: HR for stroke/SE= 0.92 (95% CI = 0.56; 1.52), HR for all cause death= 0.89 (95% CI = 0.68; 1.18). It was suggested in the European Medicines Agency (EMA) Assessment Report that the smaller treatment effect seen in the EU patients could be due to better INR control in the warfarin arm of the EU subgroup (median TTR 68.93%)[10]. This study could provide additional evidence on this point. Since all the patients in our cohort are from the UK the results can be compared with the results from the ARISTOTLE patients in the EU.

By assessing superiority and non-inferiority we will see whether the treatment effect observed is more similar to the study results as a whole (apixaban superior) or the EU subgroup results. Either a result of superiority or non-inferiority will be considered compatible with the ARISTOTLE trial results. We have set two criteria that must be met for us to conclude results are consistent with the result demonstrated in the trial:
1. The effect size must be clinically comparable with the ARISTOTLE findings; the hazard ratio for time to stroke/systemic embolism with the EHR must be between 0.69 and 0.99. This range is not symmetrical around the ARISTOTLE estimate of 0.79 as it is anticipated that the treatment effect in routine clinical care may be weaker than that seen in the optimised setting of a clinical trial.
2. The upper limit of the 95% confidence interval for the rate ratio must be less than 1.52 (upper limit in the EU subgroup of ARISTOTLE).
In addition, if the upper limit of the 95% CI is less than 1 then superiority of apixaban vs warfarin will be concluded. Either result (superiority or non-inferiority) will be taken as evidence that EHR data may be useful to look at NOACs

Secondary analyses
Secondary outcomes include the key safety outcome of major bleeding (as defined in section N) and the individual outcomes of stroke, systemic embolism, myocardial infarction, and mortality. All secondary outcomes other than major bleeding will use the same analysis approach (primary censoring scheme and regression model) as specified for the primary analysis above. For major bleeding the same regression model as the primary analysis will be used but with the supportive censoring scheme in which patients are censored around the time of last study drug. Within major bleeding results will also be summarised for intracranial, gastrointestinal, and bleeding at other locations. This safety analysis approach will be used for all 3 objectives.

Sensitivity analyses
All primary and secondary efficacy outcomes described above (stroke/SE combined and individually, MI, and mortality) will also be analysed using the supportive censoring scheme described above in which patients are censored around the time of last study drug. This analysis targets the effect of initiation of apixaban vs initiation or continuation of warfarin on the time to event while on prescribed treatment and investigates whether the extent of treatment discontinuation compromises confidence in the primary and secondary efficacy analyses.

The exclusion of patient-time post initially prescribed treatment discontinuation in the safety and sensitivity analyses might bias the results towards a conclusion of no difference[11], for example if those at higher bleeding risk were more likely to discontinue one of the treatment arms due to minor bleeding events than if those same minor

bleeding events occurred with the other treatment. The set of patients who switch or discontinue treatment during the study period will be examined to ascertain whether biases of this nature have occurred. The proportion of patients switching or discontinuing treatment, timings of withdrawal, and their baseline characteristics will be summarised by initially prescribed treatment.

Additional analyses may be performed using methods such as inverse-probability-of-censoring weighting or a rank-preserving structural failure time model to estimate the treatment effect that would have been observed in the absence of treatment switching.

We expect different adherence in routine clinical practice compared with the trial adherence may explain some of the difference in treatment effect observed between routine clinical care and the trial. Adherence will therefore be estimated in the EHR cohort to enable comparisons with the trial and investigate the extent to which this may have influenced differences in treatment effect. With the EHR data we do not know how many tablets a patient has taken or if a given prescription is filled by a patient. We will estimate the proportion of time covered by prescribing as a proxy measure for adherence; this proxy measure assumes that all prescriptions are filled and that a patient takes all tablets in the prescription. This measure of adherence is not expected to accurately estimate the adherence of a given individual but should give an idea of how adherent a patient is compared with others. Prescribing for AF in the UK is predominantly through GPs meaning prescribing information from other potential sources of treatment should not be missing. However, it is possible that a patient's first prescription may be issued in hospital and treatment prescribed during any periods of hospitalisation will not be recorded in a patient's EHR leading to missing exposure data.

We will calculate the proportion of days covered (PDC) over a patient's time when on prescribed treatment as a measure of adherence. PDC will be estimated using patient prescription data including the total number of tablets prescribed, daily dose, and number of days on treatment (derived date of last dose – index date +1). This method of estimating adherence cannot easily be used for warfarin due to daily dose being poorly recorded in EHR; in a sample of CPRD Gold warfarin prescription records many patients had dose recorded with noninformative text such as "take as directed". Depending on the quality of the prescription data warfarin adherence may instead be estimated by looking at patient adherence to other long-term daily medications as a proxy for warfarin adherence; the reliability of this measure would then be explored by comparing this proxy measure in the apixaban users with the PDC calculated directly using the apixaban prescription data.

For warfarin patients INR control will be assessed as a measure of adherence. We will look at INR values to calculate percent Time INR in Therapeutic Range (TTR) where the therapeutic range is 2.0 to 3.0 inclusive. Proportion of time in each INR interval will be calculated using Rosendaal's method. INR is influenced not only by patient adherence to the drug but also by other factors such as diet, alcohol intake, and drug interactions. The ability of a patient to comply with the lifestyle adjustments necessary to maintain INR control on a given warfarin dose may be considered as factors of warfarin treatment adherence. Patient INR control can therefore be used as a measure of overall warfarin treatment regime adherence.

We will perform a supplementary analysis comparing time to event while on treatment in patients deemed to have adequate proportion of time on therapeutic dose (adherent). We will also perform an exploratory subgroup analysis by INR TTR using TTR categories based on the TTR distribution.

P. Plan for addressing confounding

In the EHR cohort study period apixaban was a newly available treatment for the indication of interest leading to the possibility of channelling bias. The analysis cohort in this study is derived from observational data meaning there is likely to be confounding. To handle confounding for all objectives the treatment arms will be matched using the optimal method selected, for example by propensity score matching or coarsened exact matching. For objective 1 by applying the trial inclusion and exclusion criteria to both treatment cohorts and matching using the baseline covariates we should avoid channelling bias. It is possible that unmeasured or unknown confounding may remain and this will be explored and discussed in the analysis of the results.

Q. Plans for addressing missing data

Missing Baseline Data

UK EHR data have been shown to be almost complete for drug prescribing and information on important comorbidity are well recorded. The following variables used for matching may have missing data: weight and BMI, baseline systolic blood pressure, renal function, smoking status, and alcohol intake. Smoking increases the risk of stroke; we therefore anticipate that AF patients are likely to have been asked about their smoking status by their GP. A study comparing performance of ATRIA, CHADS2, and CHA2DS2-VASc risk scores in predicting stroke in patients with AF using CPRD data linked with HES found 6% of patients had smoking status not recorded[12]. Alcohol use can trigger AF symptoms so we also expect that hazardous alcohol intake will be recorded in some cases. Where there are missing data on the baseline characteristics used for matching different approaches will be taken depending on the variable in question. In some cases, such as for renal function and alcohol intake, a patient is more likely to have no data entered if there is no overt clinical evidence of abnormality; in such cases we may take a pragmatic approach and use a simplified version of the variable such as categorising into a binary parameter ("evidence of high alcohol" vs "no evidence of high alcohol intake") with those with no data included in the "no evidence of" groups. For BMI and SBP we cannot assume the data are missing at random as we expect that a patient is less likely to be weighed if they appear to be of healthy weight and is less likely to have blood pressure recorded if they do not have hypertension. Patients with missing BMI or SBP will therefore be excluded from the trial-eligible cohort. The number and proportion of patients with missing data for baseline variables will be summarised and the methods used to deal with the missing data described.

Missing Prescription Data

Treatment may be initiated in secondary care meaning the first prescription of patients newly initiating treatment is missing. To account for these potentially missing first prescriptions we will perform a sensitivity analysis where those newly initiating treatment are assumed to have a missing earlier prescription and therefore assigned an earlier derived index date. Patients who are hospitalised may also have prescriptions issued in secondary care leading to treatment gaps seen in their primary care prescription data. The primary analysis using the Prescribed Population will not be affected by such gaps as patient time is included until the patient is censored or experiences an event regardless of treatment gaps. The supplementary analysis using the On-treatment Population and the safety analysis censor at derived date of last dose and may therefore miss events whilst a patient was on treatment in hospital. We will investigate the occurrence of hospitalisation around treatment discontinuation and assess the potential impact on the results of such missed events by performing a sensitivity analysis with different extended derived dates of last dose (extending the final period to 60 or 90 days after the final prescription estimated end).

Missing Outcome Data

The problem of missing data on the outcome events will be addressed by performing a sensitivity analysis repeating the primary analysis on the ONS and HES-linked cohort. The linked cohort will be assembled by restricting the trial-eligible cohort to the patients with ONS-linked and HES-linked data prior to matching. Occurrence of events and detail of events (such as type of stroke) is expected to be better recorded in the linked cohort.

R. Patient or user group involvement (if applicable)

Patients have not been involved in the setting of this study question. However we will consult with patient groups and relevant charities in communicating the findings of the study.

S. Plans for disseminating and communicating study results, including the presence or absence of any restrictions on the extent and timing of publication

The study protocol will be submitted for publication in BMJ Open. The results of the study will be submitted to peer reviewed journals and will be presented at conferences such as the International Society of Pharmacoepidemiology conference. Results will also be published on the London School of Hygiene and Tropical Medicine website and in the PhD thesis of the principal investigator. Results that may impact on treatment guidelines will be shared with policy makers such as the Medicines and Healthcare products Regulatory Agency and the National Institute for Health and Care Excellence.

Conflict of interest statement:

The principal investigator (PI) is funded by a UK Medical Research Council PhD studentship.

T. Limitations of the study design, data sources, and analytic methods

Some of the criteria that should be assessed for ARISTOTLE eligibility may not be well recorded in CPRD. Criteria such as "increased bleeding risk" are vague and it is not clear exactly which Read codes should be included and time scale considered. Other criteria such as alcohol and drug abuse may not be captured for all patients in CPRD. These limitations are consistent with our aim to select a population as similar as possible to the ARISTOTLE trial population with the acknowledgment that differences will remain. The most important risk factors for the primary outcome of stroke (the components of the CHA2DS2-VASc score for AF stroke risk of age, sex, history of congestive heart failure, hypertension, stroke/TIA history, vascular disease history, and diabetes) are mostly well recorded in CPRD[13].

There are differences in the coding systems used by the two datasets and it is possible that the completeness of coding differs between the two. The potential impact of the different coding systems and completeness of coding will be ascertained by comparisons of the rates of diagnoses, events, baseline variables (such as smoking status, alcohol use, lab values, and vital signs), and prescriptions of interest. Including the data source (Gold or Aurum) as one of the matching variables should prevent discrepancy between the two datasets from biasing the results.

The main focus of the study is the validation of our methodology through assembling a cohort of patients comparable to the patients included in ARISTOTLE and finding similar results to the trial. Criteria to determine the success of the methodology have been pre-specified in the protocol. Given the use of CPRD data to determine treatment effectiveness is not yet well established, a finding that these data are not suitable to answer questions on intended effectiveness will be a useful conclusion.

U. References

1. Granger CB, Alexander JH, McMurray JJV, et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med. 2011 Sep 15;365(11):981-992
2. Adderley NJ, Ryan R, Nirantharakumar K, et al. Prevalence and treatment of atrial fibrillation in UK general practice from 2000 to 2016. Heart. 2019;105:27-33.
3. Davis RC, Hobbs FDR, Kenkre JE, et al. Prevalence of atrial fibrillation in the general population and in high-risk groups: the ECHOES study. Europace. 2012;14,1553-1559. doi:10.1093/europace/eus087
4. Hirsh J, Fuster V, Ansell J, Halperin JL. American Heart Association/American College of Cardiology Foundation Guide to Warfarin Therapy. 1 Apr 2003. https://doi.org/10.1161/01.CIR.0000063575.17904.4ECirculation. 2003;107:1692–1711.
5. Ruigómez A, Johansson S, Wallander MA, García Rodríguez LA, Predictors and prognosis of paroxysmal atrial fibrillation in general practice in the UK. BMC Cardiovas Disord. 2005 Jul 11;5:20
6. Iacus SM, King G, Porro G. Causal Inference without Balance Checking: Coarsened Exact Matching. Political Analysis. [Online] Cambridge University Press; 2012;20(1):1–24.
7. Berta P, Bossi M, Verzillo S. % CEM: A SAS Macro to perform Coarsened Exact Matching. J Stat Comput Simul. 2017;87(2):227-238
8. King G, Nielsen R, Coberley C, Pope JE, Wells A. 2011. Comparative Effectiveness of Matching Methods for Causal Inference. Copy at http://j.mp/2nydGlv
9. FDA Providing Regulatory Submissions in Electronic Format - Standardized Study Data: Guidance for Industry (Dec. 2014). https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM292334.pdf; https://www.cdisc.org/
10. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP). Assessment report Eliquis apixaban. Procedure No.:EMEA/H/C/002148/X/04/G. 20 September 2012. EMA/641505/2012.
11. Jones B, Jarvis P, Lewis J A, Ebbutt A F. Trials to assess equivalence: the importance of rigorous methods BMJ 1996; 313:36
12. van den Ham HA, Klungel OH, Singer DE, et al. Comparative Performance of ATRIA, CHADS2, and CHA2DS2-VASc Risk Scores Predicting Stroke in Patients With Atrial Fibrillation Results From a National Primary Care Database. J of the American College of Cardiology. 2015 Oct,66 (17):1851-1859.
13. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. Br J Gen Pract 2010; 60 (572): e128-e136. DOI: https://doi.org/10.3399/bjgp10X483562.

List of Appendices

Codelist for Atrial Fibrillation

READ code	**READ term**
14AN.00	h/o: atrial fibrillation
14AR.00	history of atrial flutter
3272.00	ecg: atrial fibrillation
3273.00	ecg: atrial flutter
G573.00	atrial fibrillation and flutter
G573000	atrial fibrillation
G573100	atrial flutter
G573200	paroxysmal atrial fibrillation
G573300	non-rheumatic atrial fibrillation
G573400	permanent atrial fibrillation
G573500	persistent atrial fibrillation
G573600	paroxysmal atrial flutter
G573z00	atrial fibrillation and flutter nos

Papers from LSHTM EHR group with outcomes or exposures used in this study

Warren-Gash, C. Herpes Zoster: Epidemiological Links With Stroke and Myocardial Infarction. J Infect Dis, 2018; 218(suppl_2):S102-S106

Silverwood, R.J. ; Forbes, H.J. ; Abuabara, K. ; Ascott, A. ; Schmidt, M. ; Schmidt, S.A.J. ; Smeeth, L. ; Langan, S.M. Severe and predominantly active atopic eczema in adulthood and long term risk of cardiovascular disease: population based cohort study. BMJ, 2018; 361:k1786

Sinnott, S.J. ; Polinski, J.M. ; Byrne, S. ; Gagne, J.J. Measuring drug exposure: concordance between defined daily dose and days' supply depended on drug class. J Clin Epidemiol, 2016; 69:107-13

S3 Supplementary Material for Paper 2

Table A1: ARISTOTLE Inclusion and Exclusion Criteria Applied to CPRD Aurum

Criteria	Implementation Rule and Notes
ARISTOTLE inclusion criteria applied to cohort	
1. Age ≥ 18 years	Day and month of birth not available therefore calculate age by assuming birthdate=01-July-birthyear.
2. Diagnosis of atrial fibrillation or atrial flutter	
3. One or more of the following risk factor(s) for stroke:	
a) Age 75 years or older	
b) Prior stroke, transient ischemic attack or systemic embolus	
c) Symptomatic congestive heart failure within 3 months or left ventricular dysfunction with an left ventricular ejection fraction (LVEF) ≤ 40%	If patient has medical record corresponding to congestive heart failure or left ventricular dysfunction diagnosis on or prior to index date.
d) Diabetes mellitus	
e) Hypertension requiring pharmacological treatment	
ARISTOTLE Exclusion criteria applied to cohort	
1. Atrial fibrillation or flutter due to reversible causes (e.g. thyrotoxicosis, pericarditis)	
2. Clinically significant (moderate or severe) mitral stenosis	Clinical significance not recorded therefore assume if there is a record of mitral stenosis condition is clinically significant.
3. Increased bleeding risk that is believed to be a contraindication to oral anticoagulation (e.g. previous intracranial hemorrhage)	
4. Conditions other than atrial fibrillation that require chronic anticoagulation (e.g. prosthetic mechanical heart valve)	
5. Persistent, uncontrolled hypertension (systolic blood pressure > 180 mm Hg, or diastolic blood pressure > 100 mm Hg)	If patient has at least 2 blood pressure readings over the limit in 6 months prior to index date OR code (within 180 days prior to index date) indicating uncontrolled hypertension.
6. Active infective endocarditis	
7. Required treatment with aspirin > 165 mg/day	
8. Simultaneous treatment with both aspirin and a thienopyridine (e.g., clopidogrel, ticlopidine)	

Criteria	Implementation Rule and Notes
9. Severe comorbid condition with life expectancy of ≤ 1 year	
10. Active alcohol or drug abuse, or psychosocial reasons that make study participation impractical	Drug or alcohol abuse or any complications of abuse, conditions involving an impaired mental state (dementia including subtypes such as Alzheimer's), severe mental health conditions (schizophrenia, psychosis, bipolar).
11. Recent ischemic stroke (within 7 days)	
12. Severe renal insufficiency (serum creatinine > 2.5 mg/dL or a calculated creatinine clearance < 25 mL/min	Lab result of serum creatinine > 2.5 mg/dL or calculated creatinine clearance < 25 mL/min within 90 days prior to index date OR code corresponding to severe renal insufficiency (chronic kidney disease stage 4 or 5, dialysis).
13. ALT or AST > 2X ULN or a Total Bilirubin ≥ 1.5X ULN (unless an alternative causative factor [e.g., Gilbert's syndrome] is identified)	Lab result showing ALT or AST > 2X ULN or a Total Bilirubin ≥ 1.5X ULN within 90 days prior to index date (AND no diagnosis of Gilbert's syndrome).
14. Platelet count ≤ 100,000/ mm^3	Lab result showing platelet count ≤ 100,000/ mm^3 within 90 days prior to index date OR a medical record of thrombocytopenia within 90 days prior to index date.
15. Hemoglobin < 9 g/dL	Lab result showing hemoglobin < 9 g/dL within 90 days prior to index date.
16. Inability to comply with INR monitoring	Evidence of drug or alcohol abuse, impaired mental state, severe mental health conditions; excluded by excl criteria number 10
17. Women of child bearing potential unwilling or unable to use an acceptable method to avoid pregnancy, women who are pregnant or breastfeeding	Exclude women with codes relating to pregnancy, childbirth, antenatal or postnatal care, or breastfeeding in the 3 years prior to index date.

ALT=alanine aminotransferase; AST=aspartate aminotransferase; CPRD=Clinical Practice Research Datalink; INR=international normalised ratio; ULN = upper limit of normal.

Note: For exclusion numbers 12-15 involving lab results a pragmatic approach was taken in which a patient was assumed not to have the exclusion criteria if there was no lab result available in the 90 days prior to index date and the latest available lab result prior to index date did not meet the criteria.

Description of modified coarsened exact matching in step 2: selection of apixaban trial-analogous patients:

A modified form of coarsened exact matching was used in which subgroups of patients were constructed based on sex, age group, prior vitamin K antagonist (VKA) exposure, $CHADS_2$ score (stroke risk factor score based on Congestive heart failure, Hypertension, Age ≥ 75 years, Diabetes, prior Stroke), stroke risk factors, and renal function category. Combining the sources of information on the ARISTOTLE patient characteristics allowed us to derive simultaneous equations relating to combinations of these subgroups which could result in a baseline distribution identical to that observed in the apixaban arm of the trial; the equations were then solved numerically giving a range of possible solutions.
Random sampling appropriate numbers of patients from these subgroups resulted in an apixaban ARISTOTLE-analogous cohort of 9,120 patients with similar baseline characteristics to ARISTOTLE participants at the point of randomisation.

Matching Feasibility

Ethnicity was limited by the pool of patients available in CPRD. We chose not to match to the trial on concomitant medications as treatment guidelines differ between countries; furthermore since oral anticoagulant (OAC) users in CPRD Aurum were matched on the stroke risk factors which are the indications for these medications, the CPRD Aurum cohort should represent typical prescribing for the trial-analogous cohort in the UK given this baseline distribution of risk factors.

Description of selection of prevalent users in Step 3: matching of apixaban trial-analogous patients to warfarin trial-eligible patients in CPRD

Continuing warfarin users in the VKA-experienced strata could be eligible for matching to the switchers to apixaban at multiple different index dates. The method for selection of index date of continuing warfarin users was not specified in the pre-published study protocol to allow testing of prevalent new user design-type methods applied to this data setting and for the objective of trial emulation. The prevalent new user design proposed by Suissa in 2017 [1] was designed to avoid the introduction of selection bias when including prevalent users; this method was unsuitable to the objective due to problems with the model convergence and complexity in constructing suitable propensity score models prior to application of the eligibility criteria.

The Webster-Clark method of sampling prevalent users was employed as detailed in Figure 1, with an adaptation to check exclusion criteria at the point of sampling, initially sampling 5 continuing warfarin users per switcher to apixaban before increasing the sample size to 10 continuing warfarin users per switcher. This method has been show in a simulation study to allow inclusion of prevalent users without introducing selection bias (Webster-Clarke 2022 [2]). The adaptation of checking exclusion criteria at the sampling stage and dropping patients from the pool of continuing warfarin users should they not be eligible at the sampled index date serves the purpose of emulating the process of screening into a randomised controlled trial (RCT).

During the sampling procedure switchers from warfarin to apixaban were taken in order of duration of prior VKA treatment history and for each switcher a sample of 10 continuing warfarin users were selected having equivalent prior VKA exposure to the switcher.

Table A2: Efficacy Outcomes Results from ARISTOTLE

	Apixaban Group (N=9,120)		Warfarin Group (N=9,081)		
Outcome	Patients with Event *no.*	Event Rate *%/yr*	Patients with Event *no.*	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	212	1.27	265	1.60	0.79 (0.66,0.95)
Stroke	199	1.19	250	1.51	0.79 (0.65,0.95)
Ischemic or uncertain type of stroke	162	0.97	175	1.05	0.92 (0.74,1.13)
Hemorrhagic stroke	40	0.24	78	0.47	0.51 (0.35,0.75)
Systemic embolism	15	0.09	17	0.10	0.87 (0.44,1.75)
Key secondary efficacy outcome: death from any cause	603	3.52	669	3.94	0.89 (0.80,0.998)
Other secondary outcomes					
Stroke, systemic embolism, or death from any cause	752	4.49	837	5.04	0.89 (0.81,0.98)
Myocardial infarction	90	0.53	102	0.61	0.88 (0.66,1.17)
Stroke, systemic embolism, myocardial infarction, or death from any cause	810	4.85	906	5.49	0.88 (0.80,0.97)
Pulmonary embolism or deep-vein thrombosis	7	0.04	9	0.05	0.78 (0.29,2.10)

CI=confidence interval ; no.=number; yr=year.
ref: C B. Granger et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2011; 365:981-992, doi: 10.1056/NEJMoa1107039 [3]

Table A3: Effectiveness Outcomes Results in the CPRD Aurum ARISTOTLE-analogous Cohort

CPRD Aurum ARISTOTLE-analogous Cohort	Apixaban Group (N=8,846)			Warfarin Group (N=8,846)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	201	15790	1.27	250	19432	1.29	0.98 (0.82,1.19)
Stroke	173	15810	1.09	216	19472	1.11	0.99 (0.81,1.21)
Ischemic or uncertain type of stroke	145	15822	0.92	157	19507	0.80	1.13 (0.90,1.41)
Hemorrhagic stroke	34	15928	0.21	65	19602	0.33	0.67 (0.44,1.01)
Systemic embolism	30	15920	0.19	35	19600	0.18	1.01 (0.61,1.66)
Key secondary efficacy outcome: death from any cause	697	15942	4.37	824	19640	4.20	1.03 (0.93,1.14)
Other secondary outcomes							
Stroke, systemic embolism, or death from any cause	846	15790	5.36	993	19432	5.11	1.04 (0.95,1.14)
Myocardial infarction	125	15837	0.79	150	19496	0.77	1.01 (0.80,1.28)
Stroke, systemic embolism, myocardial infarction, or death from any cause	934	15689	5.95	1091	19296	5.65	1.04 (0.96,1.14)
Pulmonary embolism or deep-vein thrombosis	44	15910	0.28	81	19561	0.41	0.65 (0.45,0.94)

TTR < 0.75	Apixaban Group (N=4,486)			Warfarin Group (N=4,486)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	108	7917	1.36	142	9670	1.47	0.91 (0.73,1.14)
Stroke	90	7930	1.13	122	9693	1.26	0.90 (0.70,1.14)
Ischemic or uncertain type of stroke	75	7936	0.95	91	9715	0.94	1.00 (0.76,1.32)
Hemorrhagic stroke	18	7988	0.23	35	9774	0.36	0.63 (0.38,1.04)
Systemic embolism	18	7982	0.23	21	9775	0.21	1.00 (0.55,1.83)
Key secondary efficacy outcome: death from any cause	404	7996	5.05	516	9798	5.27	0.94 (0.84,1.06)
Other secondary outcomes							
Stroke, systemic embolism, or death from any cause	483	7917	6.10	610	9670	6.31	0.95 (0.85,1.06)
Myocardial infarction	70	7937	0.88	95	9702	0.98	0.87 (0.66,1.16)
Stroke, systemic embolism, myocardial infarction, or death from any cause	532	7861	6.77	670	9580	6.99	0.95 (0.85,1.05)
Pulmonary embolism or deep-vein thrombosis	23	7979	0.29	38	9757	0.39	0.73 (0.46,1.16)

TTR ≥ 0.75	Apixaban Group (N=4,360)			Warfarin Group (N=4,360)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	91	7881	1.15	108	9761	1.11	1.05 (0.82,1.34)
Stroke	80	7890	1.01	94	9779	0.96	1.07 (0.82,1.39)
Ischemic or uncertain type of stroke	67	7896	0.85	66	9792	0.67	1.24 (0.92,1.68)
Hemorrhagic stroke	16	7944	0.20	30	9828	0.31	0.72 (0.43,1.21)
Systemic embolism	12	7942	0.15	14	9825	0.14	0.99 (0.51,1.93)
Key secondary efficacy outcome: death from any cause	298	7951	3.75	308	9842	3.13	1.20 (1.04,1.37)
Other secondary outcomes							
Stroke, systemic embolism, or death from any cause	366	7881	4.64	383	9761	3.92	1.19 (1.05,1.34)
Myocardial infarction	53	7905	0.67	55	9794	0.56	1.22 (0.88,1.70)
Stroke, systemic embolism, myocardial infarction, or death from any cause	406	7837	5.18	421	9715	4.33	1.20 (1.06,1.35)
Pulmonary embolism or deep-vein thrombosis	19	7938	0.24	43	9803	0.44	0.54 (0.35,0.84)

CI=confidence interval; CPRD=Clinical Practice Research Datalink; no.=number; TTR=time in therapeutic range; yr=year.
Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior vitamin K antagonist exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, 2.5 years after the index date).
For the analysis by TTR inverse probability of treatment weighting was applied to the apixaban users targeting the treatment effect in the warfarin users with TTR <0.75 and TTR ≥0.75.

Table A4: Bleeding Outcomes and Net Clinical Outcomes Results from ARISTOTLE RCT

ARISTOTLE RCT	Apixaban Group (N=9,088)		Warfarin Group (N=9,052)		
Outcome	Patients with Event *no.*	Event Rate *%/yr*	Patients with Event *no.*	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary safety outcome: ISTH major bleeding	327	2.13	462	3.09	0.69 (0.60,0.80)
Intracranial	52	0.33	122	0.80	0.42 (0.30,0.58)
Other location	275	1.79	340	2.27	0.79 (0.68,0.93)
Gastrointestinal	105	0.76	119	0.86	0.89 (0.70,1.15)
Net clinical outcomes					
Stroke, SE, or major bleeding	521	3.17	666	4.11	0.77 (0.69,0.86)
Stroke, SE, major bleeding, or death from any cause	1009	6.13	1168	7.20	0.85 (0.78,0.92)

CI = confidence interval; ISTH=International Society on Thrombosis and Haemostasis; no. = number; SE=systemic embolism; yr = year.

ref: C B. Granger et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2011; 365:981-992, doi: 10.1056/NEJMoa1107039 [3]

Table A5: Bleeding Outcomes and Net Clinical Outcomes Results in the CPRD Aurum ARISTOTLE-analogous Cohort

CPRD Aurum ARISTOTLE-analogous Cohort	Apixaban Group (N=8,846)			Warfarin Group (N=8,846)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary safety outcome: major bleeding	367	14998	2.45	486	17574	2.77	0.88 (0.77,1.00)
Intracranial	53	15291	0.35	89	17957	0.50	0.71 (0.51,1.00)
Other location	91	15234	0.60	114	17905	0.64	0.93 (0.70,1.22)
Gastrointestinal	230	15116	1.52	302	17717	1.70	0.88 (0.74,1.04)
Net clinical outcomes							
Stroke, SE, or major bleeding	514	14890	3.45	631	17482	3.61	0.95 (0.84,1.06)
Stroke, SE, major bleeding, or death from any cause	1005	14890	6.75	1121	17482	6.41	1.04 (0.96,1.13)

TTR < 0.75	Apixaban Group (N=4,486)			Warfarin Group (N=4,486)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary safety outcome: major bleeding	199	7489	2.66	296	8353	3.54	0.74 (0.63,0.86)
Intracranial	28	7644	0.37	51	8605	0.59	0.62 (0.41,0.92)
Other location	49	7613	0.64	75	8567	0.88	0.72 (0.52,0.99)
Gastrointestinal	127	7548	1.68	186	8447	2.20	0.75 (0.61,0.91)
Net clinical outcomes							
Stroke, SE, or major bleeding	277	7430	3.73	377	8299	4.54	0.81 (0.70,0.93)
Stroke, SE, major bleeding, or death from any cause	565	7430	7.60	677	8299	8.16	0.92 (0.83,1.02)

TTR ≥ 0.75	Apixaban Group (N=4,360)			Warfarin Group (N=4,360)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary safety outcome: major bleeding	166	7479	2.22	190	9178	2.07	1.08 (0.90,1.30)
Intracranial	24	7616	0.32	38	9304	0.41	0.80 (0.52,1.24)
Other location	42	7588	0.55	39	9290	0.42	1.35 (0.91,1.99)
Gastrointestinal	102	7538	1.35	116	9225	1.26	1.07 (0.85,1.35)
Net clinical outcomes							
Stroke, SE, or major bleeding	232	7417	3.13	254	9142	2.78	1.13 (0.97,1.32)
Stroke, SE, major bleeding, or death from any cause	440	7417	5.93	444	9142	4.86	1.22 (1.09,1.37)

CI=confidence interval; CPRD=Clinical Practice Research Datalink; no.=number; SE=systemic embolism; TTR=time in therapeutic range; yr = year.
Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, derived date of last exposure to index treatment).
For the analysis by TTR inverse probability of treatment weighting was applied to the apixaban users targeting the treatment effect in the warfarin users with TTR <0.75 and TTR ≥0.75.

Table A6: Effectiveness Outcomes Results in the CPRD Aurum ARISTOTLE-analogous Cohort using the On-treatment Censoring Scheme

	Apixaban Group (N=8,846)			Warfarin Group (N=8,846)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	196	15790	1.24	230	19432	1.18	1.04 (0.86,1.25)
Stroke	168	15810	1.06	198	19472	1.02	1.04 (0.85,1.27)
Ischemic or uncertain type of stroke	141	15822	0.89	143	19507	0.73	1.19 (0.95,1.50)
Hemorrhagic stroke	32	15928	0.20	61	19602	0.31	0.67 (0.43,1.02)
Systemic embolism	30	15920	0.19	33	19600	0.17	1.07 (0.64,1.76)
Key secondary efficacy outcome: death from any cause	662	15942	4.15	715	19640	3.64	1.12 (1.01,1.25)
Other secondary outcomes							
Stroke, systemic embolism, or death from any cause	809	15790	5.12	877	19432	4.51	1.12 (1.02,1.23)
Myocardial infarction	119	15837	0.75	133	19496	0.68	1.08 (0.84,1.38)
Stroke, systemic embolism, myocardial infarction, or death from any cause	894	15689	5.70	967	19296	5.01	1.12 (1.02,1.22)
Pulmonary embolism or deep-vein thrombosis	43	15910	0.27	73	19561	0.37	0.70 (0.48,1.02)

CI=confidence interval; CPRD=Clinical Practice Research Datalink; no.=number; yr=year.
Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, derived date of last exposure to index treatment). These results should be interpreted with caution given evidence of attrition bias in the warfarin arm.

Table A7: Treatment Status of Apixaban and Warfarin Users in CPRD Aurum ARISTOTLE-analogous Cohort during 2.5 years of Follow-up

Subject Disposition n(%) unless otherwise specified	**Apixaban** **(N=8 846)**	**Warfarin** **(N=8 846)**
Treatment persistent	7 785 (88.0)	6 805 (76.9)
On treatment until end of 2.5 year follow-up	3 120 (35.2)	5 191(58.7)
On treatment until death	591 (6.7)	607 (6.9)
On treatment until last collection date	3629 (41.0)	575 (6.5)
On treatment until registration end	445 (5.0)	432 (4.9)
Stopped treatment	519 (5.9)	596 (6.7)
Switched treatment to alternative OAC	542 (6.1)	1 445 (16.3)
Apixaban	N/A	480 (5.4)
Warfarin	149 (1.7)	N/A
Other VKA	0 (0.0)	14 (0.2)
Dabigatran	88 (1.0)	156 (1.8)
Edoxaban	94 (1.1)	60 (0.7)
Rivaroxaban	211 (2.4)	735 (8.3)
Time on treatment in months, median (IQR)	23.2 (12.2,30)	30 (20.0,30)
Time to treatment switch in months, median (IQR)	7.1 (2.9,15.2)	12.9 (5.8,21.2)

CPRD=Clinical Practice Research Datalink; IQR=interquartile range; N/A = Not applicable; n=number; OAC=oral anticoagulant; VKA=vitamin K antagonist.
Treatment persistence was ascertained using patient prescription data in CPRD Aurum with change in oral anticoagulant or gaps between prescriptions exceeding 6 months defined as distinct treatment periods.

Table A8: Characteristics of Apixaban and Warfarin Users in CPRD Aurum ARISTOTLE-analogous Cohort by Treatment Persistence During 2.5 years of Follow-up.

Characteristic	Index treatment: Apixaban			Index treatment: Warfarin		
	Apixaban persist (N=7785)	Apixaban stop (N=519)	Apixaban switch (N=542)	Warfarin persist (N=6805)	Warfarin stop (N=596)	Warfarin switch (N=1445)
Age – yr, median (IQR)	71 (63-77)	67 (57-76)	72 (63-77)	71 (64-77)	66 (57-75)	70 (62-77)
Female sex-no.(%)	2790 (35.8)	161 (31.0)	193 (35.6)	2455 (36.1)	182 (30.5)	553 (38.3)
Systolic blood pressure – mm Hg, median (IQR)	130 (120, 140)	130 (120, 140)	130 (120, 140)	130 (120, 140)	130 (120, 140)	132 (120, 140)
Weight – kg, median (IQR)	85 (73, 100)	85 (73, 99)	85 (74, 98)	85 (74, 99)	85 (74, 99)	84 (73, 99)
Prior myocardial infarction – no. (%)	964 (12.4)	52 (10.0)	74 (13.7)	833 (12.2)	63 (10.6)	178 (12.3)
Prior clinically relevant or spontaneous bleeding – no.(%)	1383 (17.8)	78 (15.0)	72 (13.3)	1142 (16.8)	103 (17.3)	262 (18.1)
History of fall within previous year – no. (%)	123 (1.6)	6 (1.2)	8 (1.5)	91 (1.3)	18 (3.0)	22 (1.5)
Prior use of vitamin K antagonist for >30 consecutive days – no. (%)	4389 (56.4)	242 (46.6)	313 (57.7)	4060 (59.7)	230 (38.6)	654 (45.3)
Qualifying risk factors						
Age ≥ 75 yr – no. (%)	2439 (31.3)	149 (28.7)	182 (33.6)	2141 (31.5)	150 (25.2)	449 (31.1)
Prior stroke, TIA, or systemic embolism – no. (%)	1537 (19.7)	73 (14.1)	101 (18.6)	1345 (19.8)	89 (14.9)	275 (19.0)
Heart failure or reduced left ventricular ejection fraction – no. (%)	2696 (34.6)	169 (32.6)	187 (34.5)	2340 (34.4)	213 (35.7)	469 (32.5)
Diabetes – no. (%)	1996 (25.6)	123 (23.7)	124 (22.9)	1779 (26.1)	138 (23.2)	358 (24.8)
Hypertension requiring treatment – no. (%)	6752 (86.7)	441 (85.0)	469 (86.5)	5908 (86.8)	508 (85.2)	1253 (86.7)
$CHADS_2$ score						
Mean ± SD	2.1 ± 1.1	1.9 ± 1.1	2.1 ± 1.1	2.1 ± 1.1	1.9 ± 1.1	2.1 ± 1.1
Distribution – no. (%)						
0	45 (0.6)	6 (1.2)	1 (0.2)	41 (0.6)	8 (1.3)	6 (0.4)
1	2552 (32.8)	221 (42.6)	198 (36.5)	2146 (31.5)	258 (43.3)	508 (35.2)

Characteristic	Index treatment: Apixaban			Index treatment: Warfarin		
	Apixaban persist (N=7785)	Apixaban stop (N=519)	Apixaban switch (N=542)	Warfarin persist (N=6805)	Warfarin stop (N=596)	Warfarin switch (N=1445)
2	2820 (36.2)	159 (30.6)	178 (32.8)	2554 (37.5)	173 (29.0)	512 (35.4)
≥3	2368 (30.4)	133 (25.6)	165 (30.4)	2064 (30.3)	157 (26.3)	419 (29.0)
Medications at index date – no. (%)						
ACE inhibitor or ARB	4931 (63.3)	265 (51.1)	333 (61.4)	4370 (64.2)	323 (54.2)	880 (60.9)
Amiodarone	295 (3.8)	20 (3.9)	21 (3.9)	238 (3.5)	25 (4.2)	59 (4.1)
Beta-blocker	5388 (69.2)	342 (65.9)	353 (65.1)	4690 (68.9)	374 (62.8)	967 (66.9)
Aspirin	440 (5.7)	37 (7.1)	37 (6.8)	413 (6.1)	49 (8.2)	95 (6.6)
Clopidogrel	204 (2.6)	11 (2.1)	14 (2.6)	162 (2.4)	11 (1.8)	42 (2.9)
Digoxin	1096 (14.1)	68 (13.1)	68 (12.5)	989 (14.5)	80 (13.4)	175 (12.1)
Calcium blocker	2650 (34.0)	148 (28.5)	167 (30.8)	2340 (34.4)	161 (27.0)	493 (34.1)
Statin	4704 (60.4)	234 (45.1)	292 (53.9)	4141 (60.9)	291 (48.8)	796 (55.1)
Nonsteroidal antinflammatory agent	429 (5.5)	28 (5.4)	30 (5.5)	345 (5.1)	35 (5.9)	99 (6.9)
Gastric antacid drugs	158 (2.0)	11 (2.1)	11 (2.0)	135 (2.0)	9 (1.5)	36 (2.5)
Proton pump inhibitor	2677 (34.4)	170 (32.8)	205 (37.8)	2342 (34.4)	210 (35.2)	552 (38.2)
H_2 receptor antagonist	249 (3.2)	12 (2.3)	20 (3.7)	189 (2.8)	11 (1.8)	50 (3.5)
Renal function, creatine clearance – no. (%)						
Normal, >80 ml/min	3586 (46.1)	276 (53.2)	236 (43.5)	3076 (45.2)	304 (51.0)	694 (48.0)
Mild impairment, >50 to 80 ml/min	2941 (37.8)	156 (30.1)	210 (38.7)	2572 (37.8)	189 (31.7)	531 (36.7)
Moderate impairment (>30 to 50 ml/min)	1116 (14.3)	74 (14.3)	86 (15.9)	1026 (15.1)	89 (14.9)	191 (13.2)
Severe impairment (le 30 ml/min)	107 (1.4)	10 (1.9)	9 (1.7)	100 (1.5)	8 (1.3)	24 (1.7)
Not reported	35 (0.4)	3 (0.6)	1 (0.2)	31 (0.5)	6 (1.0)	5 (0.3)
Other risk factors and covariates						
Peripheral artery disease – no. (%)	488 (6.3)	26 (5.0)	38 (7.0)	401 (5.9)	38 (6.4)	99 (6.9)
Aortic plaque – no. (%)	1846 (23.7)	112 (21.6)	139 (25.6)	1582 (23.2)	126 (21.1)	349 (24.2)

Characteristic	Index treatment: Apixaban			Index treatment: Warfarin		
	Apixaban persist (N=7785)	Apixaban stop (N=519)	Apixaban switch (N=542)	Warfarin persist (N=6805)	Warfarin stop (N=596)	Warfarin switch (N=1445)
Smoking status – no. (%)						
Non-smoker	2816 (36.2)	187 (36.0)	183 (33.8)	2461 (36.2)	205 (34.4)	498 (34.5)
Ex-smoker	4321 (55.5)	282 (54.3)	322 (59.4)	3801 (55.9)	322 (54.0)	822 (56.9)
Current smoker	648 (8.3)	50 (9.6)	37 (6.8)	543 (8.0)	69 (11.6)	125 (8.7)
Alcohol consumption – no. (%)						
Non-drinker	2449 (31.5)	169 (32.6)	184 (33.9)	2204 (32.4)	193 (32.4)	445 (30.8)
Light drinker, up to 14 units per week	3672 (47.2)	216 (41.6)	241 (44.5)	3202 (47.1)	269 (45.1)	672 (46.5)
Moderate drinker, 15 to 42 units per week	1366 (17.5)	105 (20.2)	92 (17.0)	1149 (16.9)	102 (17.1)	264 (18.3)
Heavy drinker, more than 42 units per week	175 (2.2)	18 (3.5)	10 (1.8)	138 (2.0)	22 (3.7)	44 (3.0)
Socioeconomic status – no. (%)						
England IMD2015 quintile 1(least deprived)	1974 (25.4)	138 (26.6)	134 (24.7)	1747 (25.7)	131 (22.0)	353 (24.4)
England IMD2015 quintile 2	1842 (23.7)	122 (23.5)	134 (24.7)	1551 (22.8)	150 (25.2)	356 (24.6)
England IMD2015 quintile 3	1509 (19.4)	88 (17.0)	118 (21.8)	1362 (20.0)	106 (17.8)	291 (20.1)
England IMD2015 quintile 4	1261 (16.2)	94 (18.1)	88 (16.2)	1108 (16.3)	105 (17.6)	252 (17.4)
England IMD2015 quintile 5(most deprived)	1199 (15.4)	77 (14.8)	68 (12.5)	1037 (15.2)	104 (17.4)	193 (13.4)
Ethnicity – no. (%)						
White	7421 (95.3)	488 (94.0)	515 (95.0)	6512 (95.7)	558 (93.6)	1374 (95.1)
Black	89 (1.1)	6 (1.2)	9 (1.7)	80 (1.2)	14 (2.3)	9 (0.6)
South Asian	175 (2.2)	14 (2.7)	15 (2.8)	134 (2.0)	16 (2.7)	41 (2.8)
East Asian	9 (0.1)	1 (0.2)	0	13 (0.2)	2 (0.3)	3 (0.2)
Mixed	19 (0.2)	4 (0.8)	2 (0.4)	19 (0.3)	2 (0.3)	7 (0.5)
Other	18 (0.2)	3 (0.6)	1 (0.2)	14 (0.2)	3 (0.5)	5 (0.3)
Unknown	40 (0.5)	2 (0.4)	0	23 (0.3)	1 (0.2)	1 (0.1)
Charlson comorbidity index components – no. (%)						

Characteristic	Index treatment: Apixaban			Index treatment: Warfarin		
	Apixaban persist (N=7785)	Apixaban stop (N=519)	Apixaban switch (N=542)	Warfarin persist (N=6805)	Warfarin stop (N=596)	Warfarin switch (N=1445)
Chronic obstructive pulmonary disease	990 (12.7)	71 (13.7)	77 (14.2)	847 (12.4)	81 (13.6)	213 (14.7)
Connective tissue disease	469 (6.0)	34 (6.6)	33 (6.1)	385 (5.7)	30 (5.0)	119 (8.2)
Peptic ulcer	356 (4.6)	19 (3.7)	36 (6.6)	284 (4.2)	26 (4.4)	83 (5.7)
Liver disease	61 (0.8)	5 (1.0)	10 (1.8)	44 (0.6)	7 (1.2)	10 (0.7)
Hemiplegia	22 (0.3)	1 (0.2)	1 (0.2)	13 (0.2)	0	3 (0.2)
Non-haematological Cancer	956 (12.3)	54 (10.4)	56 (10.3)	881 (12.9)	84 (14.1)	181 (12.5)
Haematological cancer	157 (2.0)	8 (1.5)	9 (1.7)	128 (1.9)	13 (2.2)	22 (1.5)
BMI – kg/m^2, median (IQR)	29 (26, 33)	29 (25, 33)	28 (25, 32)	29 (26, 33)	28 (25, 33)	29 (25, 33)
Time in therapeutic range, median (IQR)	N/A	N/A	N/A	0.78 (0.68, 0.86)	0.69 (0.50, 0.81)	0.64 (0.49, 0.78)

ACE = angiotensin-converting enzyme; ARB = angiotensin-receptor blocker; BMI=body mass index; $CHADS_2$=stroke risk factor score based on Congestive heart failure, Hypertension, Age ≥ 75 years, Diabetes, prior Stroke; CPRD=Clinical Practice Research Datalink; IMD2015= Index of Multiple Deprivation 2015; IQR=interquartile range; mo.=number; SD=standard deviation; TIA=transient ischemic attack; yr=year.

Treatment persistence was ascertained using patient prescription data in CPRD Aurum with change in oral anticoagulant or gaps between prescriptions exceeding 6 months defined as distinct treatment periods.

- 'persist' patients were those classified as staying on their index treatment during the 2.5 year follow-up period or until censoring
- 'stop' patients were those classified as having stopped their index oral anticoagulant treatment without evidence of any subsequent oral anticoagulant exposure in their prescription data during the follow-up period or until censoring.
- 'switch' patients were those classified as having switched from their index oral anticoagulant treatment to an alternative oral anticoagulant during the follow-up period or until censoring.

Table A9: Effectiveness Outcomes Results in the CPRD Aurum ARISTOTLE-analogous Cohort Using Later Study Start Date (01Jan2014)

All Patients	**Apixaban Group (N=8,753)**			**Warfarin Group (N=8,753)**			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	197	15667	1.26	228	19291	1.18	1.06 (0.88, 1.28)
Stroke	171	15688	1.09	199	19322	1.03	1.06 (0.88, 1.28)
Ischemic or uncertain type of stroke	144	15697	0.92	143	19351	0.74	1.23 (0.99, 1.52)
Hemorrhagic stroke	32	15805	0.20	63	19435	0.32	0.65 (0.44, 0.94)
Systemic embolism	29	15791	0.18	32	19435	0.16	1.08 (0.66, 1.77)
Key secondary efficacy outcome: death from any cause	656	15815	4.15	766	19466	3.94	1.05 (0.95, 1.16)
Other secondary outcomes							
Stroke, systemic embolism, or death from any cause	795	15667	5.07	912	19291	4.73	1.07 (0.97, 1.18)
Myocardial infarction	118	15724	0.75	129	19341	0.67	1.13 (0.87, 1.46)
Stroke, systemic embolism, myocardial infarction, or death from any cause	876	15578	5.62	998	19173	5.21	1.07 (0.97, 1.18)
Pulmonary embolism or deep-vein thrombosis	35	15787	0.22	60	19414	0.31	0.71 (0.47, 1.07)

CI=confidence interval; CPRD=Clinical Practice Research Datalink; no.=number; yr=year.

Table A10: Summary of Non-interventional Studies Comparing Apixaban and Warfarin in Atrial Fibrillation Patients

Study	Description	Results	Design differences compared to our study
Our study	CPRD Aurum linked to HES and ONS, applied trial inclusion/exclusion, matched trial on %VKA-experienced, ITT as primary analysis.	Stroke/SE 0.97 (0.83,1.13) Ischemic or uncertain stroke 1.11 (0.91,1.35) ICH 0.64 (0.46,0.89) All-cause mortality 0.99 (0.91,1.09)	
Vinogradova Y et al 2018 [4]	New users, Qresearch and CPRD Gold linked to HES and ONS, 2011-2016. Censored at treatment stop or switch. Primary prevention study. Apixaban users with AF N=10 601, Warfarin users with AF N=70 585.	Ischemic stroke 1.13 (0.89,1.44) ICH 0.40 (0.25,0.64) All-cause mortality 1.31 (1.01,1.25) Major bleeding 0.66 (0.54,0.79)	In new users alone, did not apply trial criteria or match to the trial, and excluded those with a history of the outcome event for the ischemic stroke and VTE analyses.
Larsen et al 2016 [5]	Danish databases, IPTW, standard dose apixaban only, ITT.	Ischemic stroke/SE 1.08 (0.91,1.27) All-cause mortality 0.79 (0.70,0.88)	Danish nationwide databases so may not be as applicable to UK clinical practice, used IPTW to balance covariates, did not match trial baseline characteristics, excluded patients on reduced-dose apixaban, excluded prevalent users, average follow-up 0.9 years in apixaban users. Propensity model did not include as many covariates as ours, therefore possible lower risk of death in apixaban compared with warfarin group may be caused by different baseline risk in warfarin group vs apixaban users.
Li XS, et al 2017 [6]	US claims, 1:1 PSM, after PSM 38,470 warfarin and 38,470 apixaban, 1 year follow-up, new users, some excl criteria similar to ARISTOTLE, censored at treatment switch or treatment stop+30 days	Stroke/SE 0.67 (0.59,0.76) Ischemic 0.67 (0.58,0.76) Hemorrhagic 0.70 (0.50,0.99) SE 0.46 (0.26,0.82)	In new users, in US claims data so may not be as applicable to UK clinical practice, used several of the same criteria as ARISTOTLE and had a large sample size with 38,470 PSM pairs, on-treatment analysis. 9% had $CHADS_2$=0 vs 0.6% in trial 11% on amiodarone (similar to trial which had 11% whereas we only had 4%)
Proietti et al 2018 [7]	Meta-analysis on real-world use of apixaban for stroke prevention in AF. Only 1 study from	For 'regular or any dose' subgroup: Any thromboembolic event 0.77 (0.64,0.93) Stroke 0.84 (0.69,1.01)	Systematic review and meta-analysis with only 1 small study contributing UK data.

	UK (Lee et al, n=53)	ICH 0.52 (0.44,0.61) Major bleeding 0.64 (0.51,0.80)	
J Franklin et al (protocol) [8] S V Wang et al (results) 2023 [9]	Replication of ARISTOTLE Using US Claims Data as part of the Emulation of Randomized Clinical Trials with nonrandomized Database Analyses (RCT-DUPLICATE initiative)	Stroke/SE 0.68 (0.61,0.76)	In new users alone whereas our study matches trial in proportion of VKA-experienced users, in US claims data so may not be as applicable to UK clinical practice, does not match to the trial on baseline characteristics, uses as-treated as primary analysis and ITT secondary analysis, has a shorter follow-up with maximum of 365 days, larger sample size than ours (110,259 matched pairs in protocol) and matches on a wider range of covariates.

AF=atrial fibrillation; CPRD=Clinical Practice Research Datalink; HES=Hospital Episode Statistics; CHADS2=stroke risk factor score based on Congestive heart failure, Hypertension, Age ≥ 75 years, Diabetes, prior Stroke; ICH=intracranial haemorrhage; IPTW=inverse probability of treatment weighting; ITT=intent to treat; ONS=Office of National Statistics; PSM=propensity score matched/matching; SE=systemic embolism; UK=United Kingdom; US=United States; VKA=vitamin k antagonist; VTE=venous thromboembolism.

Prediction of TTR in patients on warfarin missing TTR

To enable inclusion of patients with missing TTR in the analysis by TTR and attempt to minimise the risk of selection bias the data from the patients with TTR data was used to model TTR based on baseline variables (age, sex, BMI, smoking status, diabetes, congestive heart failure, statins, ACEi or ARB, beta-blockers, digoxin, amiodarone, NSAIDs, PPI, prior VKA exposure [naïve, <6 months prior exposure, >= 6 months prior exposure], alcohol consumption, IMD2015_5, renal function, COPD). INR values were restricted to the first year after index date to attempt to minimise selection bias. Two models were trialled for prediction: a mixed model modelling continuous TTR and a logistic regression model modelling TTR categorised as high (TTR >= 0.75) or low (TTR < 0.75) with the model successfully predicting the largest proportion of concordant pairs (observed TTR category vs predicted category) selected. The model was used to predict TTR for the patients on warfarin that were missing TTR thereby allowing all patients to be included in the analysis and attempt to minimise the risk of selection bias.

Post hoc sensitivity analysis looking at prior INR control

Prior international normalised ratio (INR) control was not included in the propensity score models for the VKA-experienced due to a high rate of missing prior INR data (missing for 34% in the apixaban arm).

An exploratory post-hoc sensitivity including a prior INR control [categorised as missing INR/poor INR control/good INR control] variable in the propensity score model for the 2 longer prior duration strata was performed. The variable on prior INR control could not be included in the shorter duration strata due to the high rate of missing data (approx. 70% missing in the apixaban shorter duration treatment strata).

Categorisation of INR control was based on the NICE criteria which specifies:

“Reassess anticoagulation for a person with poor anticoagulation control, indicated by any of the following:

- Two INR values higher than 5, or one INR value higher than 8 within the past 6 months.
- Two INR values less than 1.5 within the past 6 months
- Time in therapeutic range (TTR) is less than 65%”

Any patients meeting these NICE criteria based on their INR values in the 6 months prior to their index date were categorised as ‘poor INR control’. Patients with INR values in the 6 months prior to their index date that did not meet these criteria were categorised as ‘good INR control’ and any patients missing or insufficient INR data to determine whether they met the criteria were categorised as ‘missing INR’.

The post-hoc sensitivity analysis including this prior INR control variable in the propensity score models gave results consistent with the primary results [Stroke/SE HR 95%CI 1.02 (0.86,1.21)].

References for Appendix

[1] Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. Pharmacoepidemiol Drug Saf. 2017 Apr;26(4):459-468. doi: 10.1002/pds.4107.

[2] Webster-Clark M, Mavros P, Garry EM, Stürmer T, Shmuel S, Young J, et al. Alternative analytic and matching approaches for the prevalent new-user design: A simulation study. Pharmacoepidemiol Drug Saf. 2022 Jul;31(7):796-803. doi: 10.1002/pds.5446.

[3] C B. Granger et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. N Engl J Med 2011; 365:981-992, doi: 10.1056/NEJMoa1107039

[4] Vinogradova Y, Coupland C, Hill T, Hippisley-Cox J. Risks and benefits of direct oral anticoagulants versus warfarin in a real world setting: cohort study in primary care. BMJ. 2018 Jul 4; 362:k2505 doi:10.1136/bmj.k2505. Erratum in: BMJ. 2018 Oct 18;363:k4413.

[5] Larsen TB, Skjøth F, Nielsen PB, Kjældgaard 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 Jun 16;353:i3189. doi:10.1136/bmj.i3189.

[6] Li XS, Deitelzweig S, Keshishian A, Hamilton M, Horblyuk R, Gupta K, et al. Effectiveness and safety of apixaban versus warfarin in non-valvular atrial fibrillation patients in "real-world" clinical practice. A propensity-matched analysis of 76,940 patients. Thromb Haemost. 2017 Jun 2;117(6):1072-1082. doi:10.1160/TH17-01-0068

[7] Proietti M, Romanazzi I, Romiti GF, Farcomeni A, Lip GYH. Real-World Use of Apixaban for Stroke Prevention in Atrial Fibrillation: A Systematic Review and Meta-Analysis. Stroke. 2018 Jan;49(1):98-106. doi: 10.1161/STROKEAHA.117.018395.

[8] Franklin J, Brigham and Women's Hospital. Replication of the ARISTOTLE Anticoagulant Trial in Healthcare Claims Data. Available from: https://www.clinicaltrials.gov/ct2/show/NCT04593030

[9] Wang SV, Schneeweiss S, RCT-DUPLICATE initiative, Franklin JM, Desai RJ, Feldman W, et al. Emulation of Randomized Clinical Trials With Nonrandomized Databse Analyses: Results of 32 Clinical Trials. JAMA. 2023 Apr 25;329(16):1376-1385. doi:10.1001/jama.2023.4221.

A3.2 Additional information for Chapter 5

Table A3.2.1: Results in the CPRD Aurum ARISTOTLE-analogous cohort with minimum exposure requirement.

	Apixaban Group (N=8,734)			Warfarin Group (N=8,734)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary outcome: stroke or systemic embolism	189	15689	1.20	237	19254	1.23	0.97 (0.83-1.13)
Stroke	166	15706	1.06	209	19283	1.08	0.97 (0.82-1.14)
Ischemic or uncertain type of stroke	135	15723	0.86	149	19327	0.77	1.11 (0.91-1.35)
Hemorrhagic stroke	37	15818	0.23	70	19394	0.36	0.64 (0.46-0.89)
Systemic embolism	25	15819	0.16	31	19413	0.16	0.98 (0.60-1.61)
Key secondary efficacy outcome: death from any cause	652	15837	4.12	808	19443	4.16	0.99 (0.91-1.09)
Other secondary outcomes							
Stroke, systemic embolism, or death from any cause	795	15689	5.07	957	19254	4.97	1.02 (0.94-1.11)
Myocardial infarction	116	15735	0.74	105	19359	0.54	1.35 (1.10-1.65)
Stroke, systemic embolism, myocardial infarction, or death from any cause	877	15590	5.63	1017	19174	5.30	1.06 (0.98-1.15)
Pulmonary embolism or deep-vein thrombosis	36	15809	0.23	79	19374	0.41	0.56 (0.39-0.78)

Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, 2.5 years after the index date).

Table A3.2.2: Safety results in the CPRD Aurum ARISTOTLE-analogous cohort with minimum exposure requirement.

All Patients	Apixaban Group (N=8,734)			Warfarin Group (N=8,734)			
Outcome	Patients with Event *no.*	Person years	Event Rate *%/yr*	Patients with Event *no.*	Person years	Event Rate *%/yr*	Hazard Ratio (95% CI)
Primary safety outcome: major bleeding	364	15057	2.42	491	17691	2.78	0.85 (0.76-0.96)
Intracranial	57	15328	0.37	95	18078	0.53	0.70 (0.53-0.93)
Other location	85	15285	0.56	146	17998	0.81	0.67 (0.53-0.84)
Gastrointestinal	233	15160	1.54	272	17894	1.52	0.99 (0.86-1.14)
Net clinical outcomes							
Stroke, SE, or major bleeding	492	14959	3.29	623	17613	3.54	0.91 (0.83-1.01)
Stroke, SE, major bleeding, or death from any cause	962	14959	6.43	1084	17613	6.15	1.03 (0.96-1.11)

Note: time to event outcomes analysed using a Cox proportional hazards model with robust standard errors stratified by prior VKA exposure status. Patients were censored at the earliest of (outcome event, death, transfer out of practice, last collection date, derived date of last exposure to index treatment).

Table A3.2.3 Results in the CPRD Aurum ARISTOTLE-analogous cohort by prior VKA exposure strata

	Apixaban Event Rate %/yr	Warfarin Event Rate %/yr	Hazard Ratio (95% CI)	P value for interaction
Stroke or systemic embolism				
Prior use of VKA				0.616
No (VKA-naïve)	1.03	1.09	0.92 (0.68, 1.26)	
Yes (VKA-experienced)	1.47	1.44	1.02 (0.81, 1.29)	

	Apixaban Event Rate %/yr	Warfarin Event Rate %/yr	Hazard Ratio (95% CI)	P value for interaction
Major bleeding				
Prior use of VKA				0.417
No (VKA-naïve)	2.01	2.42	0.81 (0.65, 1.02)	
Yes (VKA-experienced)	2.79	3.04	0.91 (0.77, 1.08)	
All-cause death				
Prior use of VKA				0.914
No (VKA-naïve)	2.94	2.81	1.04 (0.86, 1.26)	
Yes (VKA-experienced)	5.49	5.34	1.03 (0.92, 1.16)	

A3.2.4 Alternative Method - Coarsened Exact Matching

For the objective of the thesis relating to methods for trial emulation the eligible new users were matched using coarsened exact matching as an alternative to propensity score matching. Coarsened exact matching (CEM) was proposed as a monotonic imbalance bounding matching method meaning the balance between the treatment groups is selected by the researcher rather than repeatedly iterating a propensity score model and checking balance with each model.

King describes the advantages of CEM as including the ability to adjust for one variable having no impact on the imbalance of any other, the user having control over the degree of model dependence, meeting the congruence principle, being robust to measurement error, and well suited to combine with multiple imputation methods for missing data.

The selection of variables on which to match is more difficult in CEM when compared with propensity score methods – whereas adding additional variables to a propensity score model will not necessarily have a large impact on the sample size, adding variables to the subgroup definition in CEM can greatly reduce the sample size. When applied to the task of RCT replication propensity score methods (assuming no stratification or variables specified for exact matching) work by balancing the included variables on aggregate whereas CEM aims to balance at the individual patient level. Thus, CEM can be seen as emulating randomisation

stratified by the variables included in the CEM subgroup definition as opposed to a simpler randomisation process emulated by propensity score methods. An advantage of CEM is that any variable included in the strata definition will define a subgroup balanced by all other variables in the strata. In propensity score methods subgroups of sufficient sample size relating to variables that were included in the propensity score model should on average be balanced across the other variables in the model, however imbalance can arise if interactions between variables have been omitted from the model or in cases of misspecification of the PS model.

As well as selecting which subset of variables to include in the CEM subgroup definition, the researcher must also decide how to combine different variables – for example whether to combine multiple binary variables into summary score measures (such as by using a comorbidity index or number of different concomitant medications). Grouping variables by means of summary indices or use of 'or' conditions will significantly increase the resulting sample size but is implicitly assuming that the components that have been combined will on average balance out between the treatment groups (ie not necessarily balance at the individual level).

To explore the suitability and performance of CEM in matching patients with AF eligible for ARISTOTLE, different selections of subgroups were trialled with the resulting sample size and balance displayed in table A7.1. Balance was assessed by means of standardised mean differences of the baseline characteristics summarised in the ARISTOTLE emulation thereby allowing a comparison with the performance of CEM against the propensity score matching of the ARISTOTLE eligible new users presented in chapter 5. Patients with NVAF eligible for OAC therapy have a range of different additional stroke factor combinations making this a key aspect to vary; three approaches for matching the stroke risk factors between patients

were trialled: firstly a 'full matching' approach requiring exact matching on all individual components of CHA_2DS_2-VASc, secondly a 'partial matching' approach in which some of the conditions were combined with 'or' logic (such as combining congestive heart failure with reduced left ventricular ejection fraction, and combining the individual vascular risk factors), and thirdly a derived risk score matching approach in which patients would be matched on the CHA_2DS_2-VASc within age and sex subgroups. The third approach would likely lead to the largest sample size but risked matching patients with different combinations of stroke risk factors at the individual level. On top of the base structure of matching the stroke risk factors, age, sex, and renal function, additional variables not balanced in the base solutions relating to comorbidities, concomitant medications, and other characteristics were added to the subgroup definitions.

Table A3.2.4 Sample size and balance using coarsened exact matching by subgroup definition for the ARITOTLE-eligible new users in CPRD Aurum

Subgroup definition	Sample size (number of matched pairs)	Balance
N/A - propensity score matching	18 684	Range [0.000, 0.032], 0.032 for aspirin, majority of variables MSD<0.01
CEM1: Age[a], sex, CHF, LVEF, hypertension, diabetes, prior stroke, prior TIA, prior SE, prior MI, PAD, aortic plaque, renal function	15 391	Perfect balance for variables in the subgroup. Other variables: 9 with MSD>=0.05, notably history of fall 0.077, ACEi or ARB 0.075, beta-blocker 0.080, aspirin 0.217, digoxin 0.071
CEM2: CEM1 + aspirin	14 357	Perfect balance for variables in the subgroup. Other variables: 8 with MSD>=0.05, notably history of fall 0.072, ACEi or ARB 0.070, beta-blocker 0.080, clopidogrel 0.061, digoxin 0.069, NSAIDs 0.055, proton pump inhibitor 0.066
CEM3: CEM1 + aspirin, fall, beta-blocker,	12, 711	Perfect balance for variables in the subgroup. Other variables: 6 with MSD>=0.05, notably ACEi or ARB 0.066, clopidogrel 0.073, digoxin 0.063, NSAIDs 0.061, proton pump inhibitor 0.061, COPD 0.054
CEM4: CEM1 + charlson[b], fall	12 800	Perfect balance for variables in the subgroup. Other variables: 8 with MSD>=0.05, notably ACEi or ARB 0.066, beta-blocker 0.078, aspirin 0.219, digoxin 0.081

CEM5: CEM1 + charlson[b], fall, medications	3 890	Perfect balance for variables in the subgroup. Other variables: 8 with MSD>=0.05, notably

a: Exact age matched for patients aged >=65 years, 5-year bins used for patients aged<65 years.
b: Modified Charlson comorbidity index including those components not excluded by the eligibility criteria and not already accounted for by the other variables included in the subgroup.
CEM = coarsened exact matching; MI = myocardial infarction; PAD = peripheral artery disease.

Table A3.2.4 shows the impact of different subgroup definitions using a 'full matching' approach for stroke risk factors requiring exact matching on all individual components of CHA_2DS_2-VASc. Different ways of combining stroke risk factors are also possible such as a 'partial matching' approach in which some of the conditions could be combined with 'or' logic (such as combining congestive heart failure with reduced left ventricular ejection fraction, and combining the individual vascular risk factors). A more relaxed exact matching could involve use of derived risk score matching in which patients would be matched on the CHA_2DS_2-VASc within age and sex subgroups. A more relaxed approach would likely lead to the largest sample size but risks matching patients with different combinations of stroke risk factors at the individual level.

Table A4.1 Baseline characteristics of the unmatched people in CPRD Aurum with AF at increased bleeding risk prescribed apixaban

Characteristic	**Apixaban (N=1853)**
Age - yr, median (IQR)	80 (73-85)
Female sex-no.(%)	776 (41.9)
Systolic blood pressure - mm Hg, median (IQR)	131 (120-140)
Weight - kg, median (IQR)	79 (68-91)
Prior myocardial infarction - no. (%)	370 (20.0)
Prior clinically relevant or spontaneous bleeding – no.(%)	835 (45.1)
History of fall within previous year – no. (%)	85 (4.6)
Prior use of vitamin K antagonist for >30 consecutive days - no. (%)	653 (35.2)
Qualifying risk factors	
Age ≥ 75 yr - no. (%)	1311 (70.8)
Prior stroke, TIA, or systemic embolism - no. (%)	732 (39.5)
Heart failure or reduced left ventricular ejection fraction - no. (%)	684 (36.9)
Diabetes - no. (%)	551 (29.7)
Hypertension requiring treatment - no. (%)	1393 (75.2)
CHADS2 score	
Mean	2.9 ± 1.4
Distribution - no. (%)	
0	77 (4.2)
1	256 (13.8)
2	473 (25.5)

Characteristic	Apixaban (N=1853)
≥3	1047 (56.5)
Medications at index date - no. (%)	
ACE inhibitor or ARB	886 (47.8)
Amiodarone	49 (2.6)
Beta-blocker	1176 (63.5)
Aspirin	93 (5.0)
Clopidogrel	75 (4.0)
Digoxin	231 (12.5)
Calcium blocker	505 (27.3)
Statin	1083 (58.4)
Nonsteroidal antinflammatory agent	119 (6.4)
Gastric antacid drugs	58 (3.1)
Proton pump inhibitor	866 (46.7)
H2 receptor antagonist	106 (5.7)
Renal function, creatine clearance - no. (%)	
Normal, >80 ml/min	437 (23.6)
Mild impairment, >50 to 80 ml/min	805 (43.4)
Moderate impairment (>30 to 50 ml/min)	546 (29.5)
Severe impairment (le 30 ml/min)	53 (2.9)
Not reported	12 (0.6)
Other risk factors and covariates	
Peripheral artery disease - no. (%)	302 (16.3)
Aortic plaque - no. (%)	658 (35.5)
Smoking status - no. (%)	
Non-smoker	648 (35.0)
Ex-smoker	1081 (58.3)
Current smoker	124 (6.7)
Alcohol consumption - no. (%)	
Non-drinker	766 (41.3)
Light drinker, up to 14 units per week	761 (41.1)
Moderate drinker, 15 to 42 units per week	197 (10.6)
Heavy drinker, more than 42 units per week	21 (1.1)
Socioeconomic status - no. (%)	
England IMD2015 quintile 1(least deprived)	488 (26.3)
England IMD2015 quintile 2	449 (24.2)
England IMD2015 quintile 3	406 (21.9)
England IMD2015 quintile 4	272 (14.7)
England IMD2015 quintile 5(most deprived)	238 (12.8)
Ethnicity - no. (%)	
White	1780 (96.1)
Black	16 (0.9)
South Asian	29 (1.6)
East Asian	5 (0.3)
Mixed	1 (0.1)
Other	6 (0.3)
Unknown	8 (0.4)
Charlson comorbidity index components - no. (%)	
Chronic obstructive pulmonary disease	363 (19.6)
Connective tissue disease	182 (9.8)
Peptic ulcer	175 (9.4)
Liver disease	21 (1.1)
Hemiplegia	12 (0.6)
Cancer	392 (21.2)
Haematological cancer	82 (4.4)
BMI - kg/m2, median (IQR)	27 (24-31)
Aneurysm or AVM	802 (43.3)
Haematuria	343 (18.5)
Gastrointestinal bleed	467 (25.2)

Characteristic	Apixaban (N=1853)
Haematological disorder	292 (15.8)
Gynaecological bleed	71 (3.8)
Prior intracranial haemorrhage	92 (5.0)
Ocular bleed	33 (1.8)
Gastrointestinal or brain tumour	36 (1.9)
Other prior bleed	192 (10.4)