

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

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

5 <sup>1</sup> Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population  
6 Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

7 <sup>2</sup> Epidemiology, GSK, London, United Kingdom

8 <sup>3</sup> Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of  
9 Hygiene and Tropical Medicine, London, United Kingdom

10 <sup>4</sup> Clinical Effectiveness Group, Centre for Primary Care, Wolfson Institute of Population Health, Barts  
11 and The London School of Medicine and Dentistry, Queen Mary University of London, London,  
12 United Kingdom

13 <sup>5</sup> Methodology and biostatistics team, Department of Efficacy and Safety, Drug sector, Saudi Food  
14 and Drug Authority, Riyadh, Saudi Arabia

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

18

19 \* Corresponding author

20 E-mail [maud.teoh@lshtm.ac.uk](mailto:maud.teoh@lshtm.ac.uk) (EMP)

**NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**

21 **ABSTRACT**

22 Background

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

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

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

34 Methods and findings

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

43 The analysis sample comprised 8734 apixaban users and propensity-score matched 8734 warfarin  
44 users in CPRD. Results [Hazard Ratio (95% Confidence Interval)] confirmed apixaban non-inferiority

45 for stroke or systemic embolism (SE) [CPRD 0.98 (0.82,1.19) vs trial 0.79 (0.66,0.95)] and death from  
46 any cause [CPRD 1.03 (0.93,1.14) vs trial 0.89 (0.80,0.998)] but did not indicate apixaban superiority.  
47 Absolute event rates for Stroke/SE were similar for apixaban in CPRD Aurum and ARISTOTLE  
48 (1.27%/year) whereas a lower event rate was observed for warfarin (CPRD Aurum 1.29%/year,  
49 ARISTOTLE 1.60%/year)

50 Analysis by TTR suggested non-inferiority of apixaban in those with TTR < 0.75 [Stroke/SE 0.94  
51 (0.75,1.19), all-cause death 0.99 (0.87,1.12)]. However, apixaban was associated with increased  
52 hazards compared with well-controlled warfarin treatment (TTR ≥ 0.75) [Stroke/SE 1.49 (1.13,1.97),  
53 all-cause death 1.75 (1.49,2.06)]. The main limitation of the study's methodology are the risk of  
54 residual confounding, channelling bias and attrition bias in the warfarin arm.

## 55 Conclusions

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

## 63 AUTHOR SUMMARY

### 64 Why Was This Study Done?

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

### 76 What Did the Researchers Do and Find?

- 77 • The researchers looked at patients with AF in a UK primary care data prescribed apixaban or  
78 warfarin and applied a “reference trial emulation” approach, in which the ARISTOTLE trial  
79 eligibility, selection and analysis approaches were applied to UK primary care data and  
80 results benchmarked against those of ARISTOTLE.
- 81 • Patients prescribed apixaban had similar rates of outcomes to those prescribed warfarin in  
82 our cohort and our results were successfully benchmarked against ARISTOTLE. Unlike  
83 ARISTOTLE we did not see superiority of apixaban vs warfarin [Hazard ratio (95% confidence  
84 interval)] for time to stroke or systemic embolism: 0.98 (0.82,1.19) in our cohort vs 0.79  
85 (0.66,0.95) in ARISTOTLE.
- 86 • We also found the benefit of apixaban vs warfarin differed depending on the quality of  
87 warfarin therapy with apixaban superior to poorly controlled warfarin therapy (TTR < 0.75)

88 [Stroke/SE 0.94 (0.75,1.19), Death 0.99 (0.87,1.12)] whereas apixaban was not superior to  
89 well-controlled warfarin (TTR  $\geq$  0.75) [Stroke/SE 1.49 (1.13,1.97), Death 1.75 (1.49,2.06)].

90 What Do These Findings Mean?

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

## 102 **Introduction**

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

109 Pharmacological therapy recommended to reduce the risk of stroke in AF includes the use of oral  
110 anticoagulants (OACs). The introduction of direct oral anticoagulants (DOACs) for AF since 2012 in  
111 the UK provided a choice of treatment alongside the older OAC class of vitamin K antagonists (VKA),

112 such as warfarin which has been available for over 60 years. The VKA OACs require regular  
113 monitoring of international normalised ratio (INR) to keep patients in the optimal therapeutic range  
114 (typically 2.0 to 3.0) in which risk of both ischemic and bleeding events are minimised [6]. A patient  
115 may require dose adjustments to stay within their INR target range. A key measure of quality of  
116 warfarin treatment is therefore the time in therapeutic range (TTR) which estimates the proportion  
117 of time a patient has spent with INR within optimal range. A TTR of 0.75 or greater is often  
118 considered as indicating optimal INR control and suggests a patient is spending a high proportion of  
119 their time in their INR target range.

120 ARISTOTLE was a pivotal RCT of the DOAC apixaban designed to demonstrate non-inferiority  
121 compared with warfarin in the prevention of stroke or systemic embolism (SE) in patients with AF.  
122 The results demonstrated superiority of apixaban over warfarin for both prevention of stroke/SE and  
123 safety (major bleeding) [7]. Results in the EU patient subset from the trial suggested the observed  
124 superiority of apixaban might be dependent on how well warfarin therapy was managed in the  
125 comparator group [8], an analysis that has not yet been performed outside of trial settings. In the  
126 NICE review of ARISTOTLE, several professional groups noted the TTR of warfarin users in ARISTOTLE  
127 may be lower than what is typical in UK clinical practice [9].

128 Treatment guidelines for DOACs are based on evidence from randomised controlled trials (RCTs),  
129 however, it is unclear whether these results extend to patient groups typically excluded from trials  
130 such as those with increased bleeding risk or severe comorbidities. Whilst there have been a number  
131 of previous studies of DOAC effectiveness using non-interventional data, there remains uncertainty  
132 on whether the data sources and methods used have fully accounted for the lack of treatment  
133 randomisation and issues such as selection bias and confounding. Comparing results from real-world  
134 studies with RCT results is challenging due to differences in patient populations, treatment  
135 adherence, and study design. However, reference trial emulation involves use of an existing named  
136 RCT to (1) inform observational study design and (2) benchmark results against, providing

137 confidence in validity of the selected observational methods and data. [10-13]. The non-  
138 interventional analysis methods can then be applied, under a set of assumptions, to reliably estimate  
139 effects in groups of patients with AF who would have been excluded from (or underrepresented in)  
140 the reference trial [14] such as patients aged > 80 that were under-represented in ARISTOTLE  
141 compared with patients with AF in UK clinical practice and patients with increased bleeding risk that  
142 were excluded by the trial eligibility criteria.

143 There is increasing interest in trial emulation using observation data, and in the application of recent  
144 developments in pharmacoepidemiology methods involving the inclusion of prevalent users. This  
145 study used a framework which involved coarsened exact matching to select patients matching the  
146 trial population on aggregate, and sampling of prevalent users in a way that avoids selection bias  
147 and emulates the process of screening into an RCT, to construct a cohort of patients similar to the  
148 target trial population which included both new and prevalent users. This methodological approach  
149 could be adapted to a variety of treatments and different therapeutic areas.

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

154

## 155 **Materials and methods**

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

## 158 **Study design**

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

## 160 **Setting/data sources**

161 UK Electronic Healthcare Records

162 This study used non-interventional data from UK Clinical Practice Research Datalink (CPRD) Aurum, a  
163 database containing anonymised data from 738 primary care practices across England  
164 (approximately 13% of the population of England with 19 million patient records and 7 million active  
165 as of September 2018 [15]. CPRD Aurum contains information on clinical diagnoses, prescribing,  
166 referrals, tests and demographic/lifestyle factors and is representative of the population of England  
167 in geographical spread, social deprivation, age and sex [15]. This study also used 2 additional data  
168 sources linked to CPRD Aurum: Hospital Episodes Statistics (HES) data, which contains data on  
169 patients admitted to NHS hospitals including diagnoses, admission and discharge, and Office of  
170 National Statistics (ONS) mortality data.

171 The reference trial (ARISTOTLE)

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

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



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

## 186 **Diagnostic and therapeutic codelists**

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

## 189 **Patient Selection**

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

191 We first selected HES-linked patients registered in CPRD Aurum between January 1, 2013 and July  
192 31, 2019, who had at least 6 months between registration and the index date. ARISTOTLE recruited  
193 both new (warfarin-naïve) and prevalent (warfarin-experienced) users of warfarin with  
194 randomisation stratified on prior warfarin (or other VKA) exposure status (warfarin naïve or  
195 experienced). To be classified as warfarin-naïve patients were required to have no evidence of  
196 exposure to warfarin or other VKA in the 5 years prior to the index date. To enable selection of a  
197 similar cohort of patients in CPRD Aurum (including both new and prevalent users of warfarin), the  
198 following process was used in determining index date:

#### 199 - apixaban users

200 index date = first prescription of apixaban in the study period  
201 apixaban user classified as warfarin-naïve or warfarin-experienced at this date;

202

#### 203 - warfarin users

204 for new users of warfarin: index date = first prescription of warfarin in the study period;  
205 for prevalent users of warfarin: a pool of potential index dates was selected containing all

206 prescription dates in the study period, with index date selected at the later treatment-history  
207 sampling stage (see step 3).

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

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

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

- 216 – Age
- 217 – Sex
- 218 – Congestive heart failure or left ventricular systolic dysfunction
- 219 – Hypertension requiring treatment
- 220 – Diabetes mellitus
- 221 – Prior stroke/transient ischaemic attack (TIA)/systemic embolism (SE)
- 222 – Level of renal impairment
- 223 – Prior VKA/warfarin exposure

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

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

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

230 To emulate ARISTOTLE which stratified randomisation on prior VKA exposure status, patients in the  
231 CPRD cohort were matched separately within the VKA-naïve and VKA-experienced strata. A 3-step  
232 procedure, based on methods proposed by Suissa et al [20] and Webster Clark et al [21], was used to  
233 select and match patients in the VKA-experienced strata whilst avoiding selection bias; this  
234 procedure is summarised in Figure 1 and described in in S3 Appendix.

235

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

240

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

242 **Table 1: Covariates Included in the Propensity Score Models**

Category	Variable List
Demographics	age, sex, ethnicity
CHADS <sub>2</sub> 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

243 AF=atrial fibrillation; COPD=chronic obstructive pulmonary disease; GP=general practitioner;  
244 NSAIDs=non-steroidal anti-inflammatory drugs; TIA=transient ischemic attack;  
245

246 The model resulting in the most balanced cohort was chosen with balance assessed by looking at  
247 standardised differences across all variables after matching using a target threshold of 0.05 for the  
248 maximum difference allowed for any individual variable. Balance of covariates considered to be  
249 most important in predicting outcome were prioritised namely age, sex, and stroke risk factors.

250

251  
252  
253  
254  
255  
256

**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.

257 **Exposures and outcomes**

258 Exposures

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

261 Outcomes

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

271 Statistical analysis

272 **Methods of Analysis**

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

275 Changes from the planned protocol are described in Table 2.

276 **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 $\geq$ 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.

277

278 All time-to-event endpoints were analysed using a Cox proportional hazards model, stratified by

279 prior VKA status (experienced, naïve). The effectiveness outcomes were analysed using the

280 intention-to-treat principle and major bleeding was analysed using an on-treatment censoring

281 scheme. Patients were censored at 2.5 years after index date reflecting typical maximum duration of

282 follow-up in ARISTOTLE. Cluster-robust standard errors were used with pair membership as the

283 clustering variable [24,25]. The proportional hazards assumption was assessed by looking at the log-

284 log of the Kaplan-Meier survival curves and inspection of scaled Schoenfeld residuals plotted against  
285 time. Analyses were performed using SAS version 9.4 and R version 4.2.1.

## 286 **Supplementary analyses**

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

297 An additional post hoc analysis was performed looking at the proportion of apixaban patients  
298 prescribed reduced-dose apixaban along with a comparison of the patients meeting the criteria for  
299 dose-reduction against the dose actually prescribed. In this analysis apixaban dose in the ARISTOTLE-  
300 analogous CPRD cohort was assessed and compared against the ARISTOTLE protocol-specified  
301 criteria and NICE criteria for reduced apixaban dose. ARISTOTLE specified that participants meeting  
302 any 2 of the following criteria assessed at the time of randomisation should have their apixaban dose  
303 reduced to 2.5mg BID: age  $\geq 80$  years, body weight  $\leq 60$  kg, or serum creatinine  $\geq 1.5$  mg/dL. These  
304 criteria are equivalent to the NICE guidelines for dose reduction with NICE having an additional  
305 criteria indicating reduced dose in those with creatinine clearance 15–29 mL/minute.

306 In addition, to assess the impact of the quality of dose-adjustment in the CPRD cohort on the  
307 observed effectiveness of apixaban relative to warfarin, a supplementary post hoc analysis was



308 performed looking at the results in the subset of apixaban patients prescribed the correct dose  
309 compared with IPTW re-balanced warfarin comparators.

### 310 **Sensitivity analyses**

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

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

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

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

331 **Confounding and bias**

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

336 **Benchmarking results against ARISTOTLE**

337 The study hypothesis was that results in the CPRD ARISTOTLE-analogous cohort would be  
338 comparable to the ARISTOTLE results, as defined by the pre-specified benchmarking criteria. A  
339 slightly weaker benefit of apixaban vs warfarin was expected based on the weaker benefit seen in  
340 the EU subgroup of ARISTOTLE and an expectation that the quality of warfarin control in UK patients  
341 may be higher than that observed in ARISTOTLE.

342 The benchmarking criteria for considering the results in the trial-analogous CPRD cohort to be  
343 comparable with ARISTOTLE were pre-specified and published previously [14]:

- 344 1. The effect size must be clinically comparable with the ARISTOTLE findings; the HR for time to  
345 stroke/SE with the HR must be between 0.69 and 0.99. This range is not symmetrical around  
346 the ARISTOTLE estimate of 0.79 as it is anticipated the treatment effect in routine clinical  
347 care may be weaker than that seen in the optimised setting of a clinical trial.
- 348
- 349 2. The upper limit of the 95% CI for the HR for time to stroke/SE must be less than 1.38 (non-  
350 inferiority margin used in ARISTOTLE, updated since protocol – see Table 2).

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

## 356 **Missing data**

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

## 365 **Ethics**

366 Scientific approval was provided by the London School of Hygiene and Tropical Medicine research  
367 ethics committee (ref 17682) and the independent scientific advisory committee of the Medicines  
368 and Healthcare Products Regulatory Agency (protocol no. 19\_066R). CPRD data are already  
369 approved via a national research ethics committee for purely non-interventional research of this  
370 type. CPRD data are analysed anonymously therefore individual patient consent is not sought by  
371 contributing medical practices when data is shared with CPRD; however, patients are able to opt out  
372 of their patient information being shared for research.

## 373 **Results**

### 374 **Participants**

375 Between January 1, 2013 and July 31, 2019 there were 86,888 people with AF prescribed apixaban  
376 and 159,632 prescribed warfarin in HES-linked CPRD Aurum practices (Figure 2). Application of  
377 minimum registration period and ARISTOTLE inclusion criteria reduced this to 67,539 apixaban and  
378 139,527 warfarin patients. After applying ARISTOTLE exclusion criteria there were 41,487 apixaban  
379 and 101,159 warfarin patients.

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

384

### 385 **Figure 2: Selection of ARISTOTLE-analogous CPRD Aurum Cohort**

386 Flow of number of individuals included in the analysis. AF = atrial fibrillation; ALT = alanine  
387 transaminase; AST = aspartate transaminase; BMI = body mass index; BP = blood pressure; CPRD =  
388 Clinical Practice Research Datalink; HES: Hospital Episodes Statistics; Rx = Prescription; SES =  
389 socioeconomic status; ULN = upper limit of normal; VKA = vitamin K antagonist.

390 a Severe comorbid condition with life expectancy <1 year or reasons making participation  
391 impractical; b ALT or AST > 2X ULN or Total Bilirubin  $\geq$  1.5X ULN; c Pregnant or breastfeeding within 3  
392 years prior

393 See supplementary table A1 in S3 Appendix for detailed list of inclusion and exclusion criteria.

394 Note: For prevalent warfarin users trial eligibility only revealed at point of random selection into the  
395 cohort for prevalent users. Numbers in figure show maximum theoretical number of warfarin users  
396 available should they be selected only at a time they were eligible for the trial.

397

### 398 **Application of ARISTOTLE inclusion/exclusion criteria and matching to ARISTOTLE**

399 Applying the ARISTOTLE inclusion/exclusion criteria and matching to ARISTOTLE baseline patient  
400 characteristics resulted in a cohort similar to the ARISTOTLE apixaban participants (Table 3); for  
401 example median age was 78 and mean CHADS<sub>2</sub> score 2.4 in CPRD Aurum before applying trial criteria  
402 and matching whereas the median age of 71 and mean CHADS<sub>2</sub> score 2.1 after these steps matched  
403 the ARISTOTLE apixaban participants. The ARISTOTLE-analogous apixaban arm matched the trial arm  
404 on prior VKA exposure, age, sex, stroke risk factors and CHADS<sub>2</sub> score, and proportion of patients  
405 with moderate or severe renal impairment.

406 Differences remained on baseline characteristics it was not feasible to match on namely: ethnicity  
407 (95.2% white, 2.4% Asian in CPRD Aurum apixaban vs 82.6% white, 14.4% Asian in ARISTOTLE) and  
408 concomitant medications (amiodarone 3.8%, aspirin 5.8%, digoxin 13.9% in CPRD Aurum apixaban  
409 users vs amiodarone 11.1%, aspirin 31.3%, digoxin 32.0% in ARISTOTLE apixaban arm). See Appendix  
410 for details on matching feasibility.

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

413 **Results of Propensity score matching**

414 Before propensity score matching, differences between treatment groups were evident for most  
415 baseline variables including age (median age 71 in apixaban vs 78 in warfarin), sex (apixaban 35.6%  
416 female vs warfarin 43.6%), and stroke risk factors [see Table 3]. After propensity score matching all  
417 baseline characteristics were well balanced (maximum standardised difference .031). From 9120  
418 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)

<b>CHADS<sub>2</sub> score, mean ± SD</b>	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
<b>CHADS<sub>2</sub>=0</b>	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)
<b>CHADS<sub>2</sub>=1</b>	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)
<b>CHADS<sub>2</sub>=2</b>	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)
<b>CHADS<sub>2</sub> ≥3</b>	32 645 (44.2)	64 674 (44.2)	17 783 (42.9)	44 284 (43.8)	2666 (30.1)	2640 (29.8)	0.006	2 758 (30.2)	2 744 (30.2)
<b>Medications at index date</b>									
<b>ACE inhibitor or ARB</b>	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)
<b>Amiodarone</b>	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)
<b>Beta-blocker</b>	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)
<b>Aspirin</b>	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)
<b>Clopidogrel</b>	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)
<b>Digoxin</b>	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)
<b>Calcium channel blocker</b>	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)
<b>Statin</b>	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)
<b>Non-steroidal anti-inflammatory</b>	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)
<b>Gastric antacid drugs</b>	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)
<b>Proton pump inhibitor</b>	2844 (38.0)	47 838 (32.7)	15 197 (36.6)	31 769 (31.4)	3052 (34.5)	3104 (35.1)	0.012		
<b>H2 receptor antagonist</b>	3 188 (4.3)	4 837 (3.3)	1 586 (3.8)	3 006 (3.0)	281 (3.2)	250 (2.8)	0.021		

<b>Renal function, creatinine clearance</b>									
<b>Normal, &gt;80 ml/min</b>	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)
<b>Mild imp., &gt;50 to 80 ml/min</b>	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)
<b>Moderate imp. (&gt;30 to 50 ml/min)</b>	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)
<b>Severe imp. (≤30 ml/min)</b>	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)
<b>Not reported</b>	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)
<b>Peripheral artery disease</b>	5 984 (8.1)	12 764 (8.7)	2 770 (6.7)	7 516 (7.4)	552 (6.2)	538 (6.1)	0.007		
<b>Aortic plaque</b>	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		
<b>Smoking status</b>									
<b>Non-smoker</b>	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		
<b>Ex-smoker</b>	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		
<b>Current smoker</b>	5 236 (7.1)	9 658 (6.6)	2 688 (6.5)	6 049 (6.0)	735 (8.3)	737 (8.3)	0.001		
<b>Missing</b>	224	211	94	102	0	0			
<b>Alcohol consumption</b>									
<b>Non-drinker</b>	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		
<b>Light, 1 to 14 units per week</b>	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		
<b>Moderate, 15 to 42 units per week</b>	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		
<b>Heavy, &gt; 42 units per week</b>	1 488 (2.0)	2 028 (1.4)	617 (1.5)	1 149 (1.1)	203 (2.3)	204 (2.3)	0.001		



<b>Missing</b>	3 901	9 223	2 032	5 893	143	132			
<b>Socioeconomic status (England IMD2015 Quintile)</b>									
<b>1 (least deprived)</b>	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		
<b>2</b>	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		
<b>3</b>	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		
<b>4</b>	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		
<b>5 (most deprived)</b>	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		
<b>Missing</b>	69	165	36	116	0	0			
<b>Ethnicity</b>									
<b>White</b>	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)
<b>Black</b>	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)
<b>Asian</b>	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)
<b>Other</b>	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)
<b>Mixed</b>	152 (0.2)	308 (0.2)	75 (0.2)	190 (0.2)	25 (0.3)	28 (0.3)	0.006	0	0
<b>Unknown</b>	385 (0.5)	448 (0.3)	252 (0.6)	350 (0.3)	42 (0.5)	25 (0.3)	0.031	0	0
<b>Charlson comorbidity components</b>									
<b>Chronic obstructive pulmonary disease</b>	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		
<b>Connective tissue disease</b>	5 377 (7.3)	9 784 (6.7)	3 000 (7.2)	6 744 (6.7)	536 (6.1)	534 (6.0)	0.001		

<b>Peptic ulcer</b>	4 400 (6.0)	8 399 (5.7)	2 161 (5.2)	5 458 (5.4)	411 (4.6)	393 (4.4)	0.010		
<b>Liver disease</b>	761 (1.0)	1 291 (0.9)	263 (0.6)	642 (0.6)	76 (0.9)	61 (0.7)	0.019		
<b>Hemiplegia</b>	265 (0.4)	559 (0.4)	147 (0.4)	379 (0.4)	24 (0.3)	16 (0.2)	0.019		
<b>Non-haematological Cancer</b>	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		
<b>Haematological cancer</b>	1 966 (2.7)	3 481 (2.4)	951 (2.3)	2 231 (2.2)	174 (2.0)	163 (1.8)	0.009		
<b>BMI - kg/m<sup>2</sup>, median (IQR)</b>	28 (24, 32)	28 (23, 32)	28 (25, 32)	28 (25, 32)	29 (26, 33)	29 (26, 33)	0.003		
<b>Missing</b>	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<sub>2</sub> = 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;

## 419 Main results

420 The hazard ratio (HR) for stroke/systemic embolism (SE) in the propensity score matched groups was  
421 0.98 (95% CI 0.82,1.19) (Figure 3 and Table A3 in S3 Appendix). This association was consistent with  
422 the non-inferiority margin (upper limit of the 95% CI less than 1.38) [7] but did not show superiority  
423 as predicted by ARISTOTLE [HR 0.79 (95% CI 0.66,0.95)] (Figure 3 and Appendix Table A2 in S3  
424 Appendix). The outcome of all-cause mortality also showed non-inferiority [Aurum 1.03 (0.93,1.14)  
425 vs trial 0.89 (0.80,0.998)] but did not indicate apixaban superiority. Absolute event rates for the  
426 primary outcome and components were close to the trial for apixaban – for example [comparing  
427 Aurum vs trial] stroke/SE event rate of 1.27%/year vs. 1.27% whereas the warfarin group had a  
428 lower event rate compared with ARISTOTLE (stroke/SE event rate of 1.29%/year vs. 1.60% and  
429 hemorrhagic stroke 0.33 %/yr vs 0.47%/yr) (Figure 3). Mean duration of follow-up in the cohort was  
430 1.8 years in the apixaban arm and 2.2 years in the warfarin arm.

431 **Figure 3: Forest plot showing hazard ratios (dots) and 95% confidence intervals (lines) for**  
432 **apixaban vs warfarin. Absolute event rates (%/year) and Hazard Ratio (95% Confidence**  
433 **Intervals) are presented for key effectiveness outcomes in i) ARISTOTLE, ii) CPRD Aurum**  
434 **trial-matched cohort, iii) CPRD Aurum trial-matched with TTR<0.75, and iv) CPRD Aurum**  
435 **trial-matched with TTR≥0.75.**

436 **Dashed line shows non-inferiority margin 1.38 for the upper bound of the 95% CI of the**  
437 **hazard ratio used in ARISTOTLE for the primary outcome of stroke or systemic embolism.**  
438 **For the analysis by TTR inverse probability of treatment weighting was applied to the**  
439 **apixaban users targeting the treatment effect in the warfarin users with TTR <0.75 and**  
440 **TTR ≥0.75.**

441 **CI=confidence interval; CPRD=Clinical Practice Research Datalink; HR=hazard ratio;**  
442 **TTR=time in therapeutic range.**  
443

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

445 TTR was higher in the CPRD cohort than in ARISTOTLE (mean 0.73 vs. 0.62, median 0.76 vs 0.66).

446 Analysis by TTR suggested non-inferiority of apixaban vs warfarin in those with TTR < 0.75 [Stroke/SE  
447 0.94 (0.75,1.19), all-cause death 0.99 (0.87,1.12)] (Figure 3). Apixaban was associated with increased

448 hazards of both outcomes compared to warfarin in those with well-controlled warfarin treatment  
449 (TTR  $\geq$  0.75) [Stroke/SE 1.49 (1.13,1.97), all-cause death 1.75 (1.49,2.06)] (Figure 3).

450 **Analysis of apixaban dose-adjustment**

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

458 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

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

461 A further analysis of the quality of dose-adjustment in patients in CPRD Aurum indicated 10.5% of  
462 patients may have been prescribed an incorrect dose of apixaban at the index prescription based on  
463 the data contained in their EHRs. The majority of incorrect dose relating to patients being prescribed  
464 reduced-dose apixaban despite not meeting the criteria for dose reduction. A large proportion of  
465 patients prescribed an incorrect dose had only 1 dose adjustment criteria (59.6% of those with  
466 incorrect dose) suggesting some prescribers may have thought a dose reduction was warranted

467 when only 1 criteria was present. Other possible reasons for the incorrect dose-adjustment observed  
 468 here may be data on the criteria missing from the EHR record (ie incorrect ascertainment) or  
 469 consideration of other medical history which made a prescriber adjust the dose.

470 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%)

471  
 472 To assess the impact of the quality of dose-adjustment in the CPRD cohort on the effectiveness of  
 473 apixaban a supplementary post hoc analysis was performed looking at the results in the subset of  
 474 apixaban patients prescribed the correct dose (N=7921) compared with IPTW re-balanced warfarin  
 475 comparators. The results in this subset were consistent with the primary results showing apixaban to  
 476 be non-inferior to warfarin (Stroke/SE 0.96 [0.78,1.17], death 0.97 [0.87,1.09]) with the results  
 477 moving slightly closer to those observed in ARISTOTLE.

478 **Safety results**

479 The analysis for safety outcomes is presented in Figure 4 and Table A5 in S3 Appendix; patients on  
 480 apixaban had a lower risk of major bleeding compared with those on warfarin, HR (95% CI) 0.88  
 481 (0.77,1.00), consistent with ARISTOTLE. Analysis by TTR suggested superiority of apixaban for major  
 482 bleeding in those with TTR <0.75 [0.70 (0.59,0.82)] whereas apixaban users had a higher risk of  
 483 major bleeding compared with those with optimal warfarin control (TTR ≥ 0.75) [1.31 (1.08,1.59)].

484

485 **Figure 4: Forest plot showing hazard ratios (dots) and 95% confidence intervals (lines) for**  
486 **apixaban vs warfarin. Absolute event rates (%/year) and Hazard Ratio (95% Confidence**  
487 **Intervals) are presented for key safety outcomes in i) ARISTOTLE, ii) CPRD Aurum trial-**  
488 **matched cohort, iii) CPRD Aurum trial-matched with TTR<0.75, and iv) CPRD Aurum trial-**  
489 **matched with TTR≥0.75.**

490 **For the analysis by TTR inverse probability of treatment weighting was applied to the**  
491 **apixaban users targeting the treatment effect in the warfarin users with TTR <0.75 and**  
492 **TTR ≥0.75.**

493 **CI=confidence interval; CPRD=Clinical Practice Research Datalink; HR=hazard ratio;**  
494 **TTR=time in therapeutic range.**

495

496 **Sensitivity analyses**

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

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

506 On-treatment analyses censoring around treatment switch or discontinuation are presented for the  
507 effectiveness analyses in the appendix (Table A6 in S3 Appendix); the results show evidence of the  
508 expected attrition bias against apixaban when compared with the ITT results in Figure 2, for example  
509 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  
510 the ITT analysis.

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

513 Prior INR control was not included in the propensity score models for the VKA-experienced due to a  
514 high rate of missing prior INR data (missing for 34% in the apixaban arm). A post-hoc sensitivity

515 analysis including a prior INR control variable in the PSM gave results consistent with the primary  
516 results [Stroke/SE HR 95%CI 1.02 (0.86,1.21)]. Details of this post hoc analysis are in S3 Appendix.

## 517 **Discussion**

518 In our emulation of ARISTOTLE using UK routinely-collected healthcare data we found results that  
519 met our predefined criteria for comparability with the trial. We saw non-inferiority of apixaban vs  
520 warfarin for prevention of stroke or systemic embolism, all-cause mortality, and major bleeding, but  
521 did not see superiority of apixaban vs warfarin for these outcomes as was seen in ARISTOTLE. We  
522 found higher TTR in the patients using warfarin in our cohort compared with the warfarin arm of  
523 ARISTOTLE (median 0.76 vs 0.66). Our analysis by TTR showed that in patients well-controlled on  
524 warfarin (TTR  $\geq$  0.75) warfarin is superior to apixaban whereas in patients not well-controlled on  
525 warfarin (TTR < 0.75) apixaban is non-inferior. We saw evidence suggesting sub-optimal dosing of  
526 apixaban in our cohort with approximately 10% of patients in the apixaban arm prescribed the reduced  
527 dose without meeting the criteria for the reduced dose.

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

532 Our findings are consistent with a UK study of ischemic stroke which compared DOACs with warfarin  
533 [29]. A Danish study found similar results to ours for stroke/SE [30] although they found apixaban  
534 users had a lower risk of death, a study of US claims data [31] also found apixaban was associated  
535 with a lower risk of death. A systematic review and meta-analysis of 16 studies [32] found pooled  
536 results for stroke and ICH that were consistent with ours. One study (in US claims data) also aimed  
537 to replicate ARISTOTLE [33, 34] and in contrast to our study found superiority for apixaban for  
538 stroke/SE, which may be linked to population differences such as lower TTR in US patients on  
539 warfarin [35] and differences in ethnicity. None of these studies matched to the ARISTOTLE trial

540 participants, included prevalent users, or looked at how warfarin control impacted results. Further  
541 details on these studies including design and key results are summarized in Table A10 in S3  
542 Appendix.

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

557 An additional strength of our study was the ability to explore the quality of warfarin treatment in our  
558 cohort and the impact of INR control on the treatment effect estimates. Our finding that the benefits  
559 of apixaban vs warfarin depended on the quality of INR control in the warfarin arm answers  
560 questions raised in the NICE premeeting briefing which looked at apixaban in the NVAF population  
561 and noted the TTR seen in ARISTOTLE “may be lower than what is typical in UK clinical practice” and  
562 “apixaban compared with well-controlled warfarin (TTR 75% or more) may not be superior in the  
563 long term” [8]. ARISTOTLE presented outcomes by centre (for example hospital) TTR quartile and did  
564 not show a signal of treatment efficacy differing by centre TTR quartile. We were able to use inverse



565 probability of treatment weighting to estimate the treatment effect in the different warfarin TTR  
566 groups. Our findings may help inform choice of treatment in the context of prior warfarin treatment  
567 and availability of INR measurements.

568 Whilst our study aimed to emulate ARISTOTLE using suitable methods there were several limitations.  
569 Some of the criteria assessed for ARISTOTLE eligibility may not be well recorded in CPRD leading to a  
570 risk of misclassification. Furthermore, misclassification of ARISTOTLE eligibility criteria and baseline  
571 covariates could be differential by treatment in the VKA-experienced patients if criteria such as renal  
572 function are more likely to be checked before changing treatment. However, the most important risk  
573 factors for the primary outcome of stroke (the components of CHA<sub>2</sub>DS<sub>2</sub>-VASc stroke risk score) are  
574 mostly well recorded in CPRD Aurum and HES.

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

583 A key limitation of our study was the inability to match ARISOTLE on ethnicity meaning the CPRD  
584 Aurum cohort included a low number of patients from Asian and Hispanic groups when compared  
585 with the RCT (14.5% of participants in ARISTOTLE were Asian compared to 2.4% in our ARISTOTLE-  
586 analogous CPRD cohort). There are known racial differences in the treatment effects of OACs with  
587 Asian patients experiencing a higher risk of haemorrhagic stroke and intracranial haemorrhage  
588 compared with White patients; in ARISTOTLE Asian participants experienced double the risk of  
589 stroke or systemic embolism when on warfarin therapy when compared with White participants

590 [37]. The reasons for the increased risk of bleeding associated with warfarin therapy in Asian  
591 patients is hypothesised to be associated with differences in drug metabolism and prevalence of  
592 cerebral microbleeds [38]. The difference in proportion of Asian patients between our cohort and  
593 ARISTOTLE is therefore likely to explain some of the differences in treatment effects seen and limits  
594 the generalisability of our study, with the results of our study of most relevance to White patients.  
595 This limitation on ethnicity arose from the data source used and time period studied (patients with  
596 AF in CPRD Aurum 2013-2019) which had a low proportion of Asian patients, likely due to AF being  
597 associated with older age combined with a lower prevalence of AF in Asian patients compared with  
598 White patients [38]. Whilst CPRD Aurum is largely representative of the UK population in relation to  
599 ethnicity [39], diversity is still limited for older individuals. Despite this, CPRD Aurum has shown to  
600 be a useful resource for investigating treatment effects in different ethnic groups for indications  
601 such as hypertension which is more prevalent and occurs at a younger age in ethnic minority groups,  
602 with similar trial replication methods used to compare antihypertensive treatment effects in  
603 underrepresented ethnic groups [40].

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

616 across multiple outcomes providing reassurance the method used was likely to have successfully  
617 avoided selection bias.

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

629 Adherence to treatment was difficult to assess in our study due to automatic repeat prescriptions;  
630 treatment persistence was more useful in providing a measure of pattern of medicine use over time.  
631 In the analysis by TTR the adherence of patients using apixaban was not accounted for, however, a  
632 previous UK study showed apixaban had higher adherence than VKAs [41] meaning we would expect  
633 to see better effectiveness outcomes in apixaban. Furthermore, the use of IPTW in the analysis by  
634 TTR means predictors of poor adherence are likely to have been balanced between treatments. The  
635 analysis of TTR is limited by this being a post-baseline measure available for only one treatment arm;  
636 this limitation was also evident in the RCTs of DOACs vs warfarin and is mitigated in our study  
637 through the use of IPTW. Sensitivity analyses in our cohort using an on-treatment censoring scheme  
638 showed evidence of attrition bias. The regular measurement of INR and availability of alternative  
639 anticoagulants makes warfarin therapy particularly prone to attrition bias since a patient may be

640 more likely to switch to a DOAC if their INR is frequently out of the optimal range or if they have not  
641 been adhering to scheduled INR testing.

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

653 We further showed that when compared to well-controlled warfarin therapy, apixaban was  
654 associated with harmful treatment effects. The weaker overall treatment benefit observed in our  
655 cohort appears to be due to a higher proportion of patients with well-controlled warfarin in the UK  
656 clinical context, compared with the trial. NICE guidance (2021) advises that for patients already on  
657 VKA their TTR should be taken into account when deciding whether to switch to a DOAC; our  
658 findings support the option of staying on VKA for those patients happy with their treatment and who  
659 have  $TTR \geq 0.75$ .

660 S1: STROBE checklist

661

662 S2: ISAC protocol for the ARISTOTLE emulation study

663

664 S3: Appendix containing supporting information

665

666 **References**

- 667 [1] Adderley NJ, Ryan R, Nirantharakumar K, Marshall T. Prevalence and treatment of atrial  
668 fibrillation in UK general practice from 2000 to 2016. *Heart*. 2019 Jan; 105(1):27-33. doi:  
669 10.1136/heartjnl-2018-312977.
- 670 [2] Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the  
671 Framingham Study. *Stroke*. 1991 Aug;22(8):983-8. doi: 10.1161/01.str.22.8.983.
- 672 [3] Friberg L, Rosenqvist M, Lindgren A, Terént A, Norrving B, Asplund K. High prevalence of atrial  
673 fibrillation among patients with ischemic stroke. *Stroke*. 2014 Sep;45(9):2599–605. doi:  
674 10.1161/STROKEAHA.114.006070.
- 675 [4] Ali AN, Abdelhafiz A. Clinical and Economic Implications of AF Related Stroke. *J Atr Fibrillation*.  
676 2016 Feb. 29;8(5):1279. doi: 10.4022/jafib.1279.
- 677 [5] Andrew NE, Thrift AG, Cadilhac DA. The prevalence, impact and economic implications of atrial  
678 fibrillation in stroke: what progress has been made? *Neuroepidemiology*. 2013; 40(4):227–39. doi:  
679 10.1159/000343667.
- 680 [6] Apostolakis S, Sullivan RM, Olshansky B, Lip GYH. 053 Factors affecting quality of warfarin  
681 anticoagulation in patients with atrial fibrillation: insights from affirm. *Heart*. 2012;98:A32.  
682 doi:10.1136/heartjnl-2012-301877b.53.
- 683 [7] Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, et al. Apixaban versus  
684 warfarin in patients with atrial fibrillation. *N Engl J Med*. 2011 Sep 15;365(11):981-92, doi:  
685 10.1056/NEJMoa1107039.
- 686 [8] European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP)  
687 Assessment report Eliquis apixaban. Procedure No.:EMA/H/C/002148/X/04/G. EMA/641505/2012,  
688 2012, page 35. Available from: [https://www.ema.europa.eu/en/documents/variation-report/eliquis-](https://www.ema.europa.eu/en/documents/variation-report/eliquis-h-c-2148-x-0004-g-epar-assessment-report-extension_en.pdf)  
689 [h-c-2148-x-0004-g-epar-assessment-report-extension\\_en.pdf](https://www.ema.europa.eu/en/documents/variation-report/eliquis-h-c-2148-x-0004-g-epar-assessment-report-extension_en.pdf)
- 690 [9] National Institute for Health and Clinical Excellence. Premeeting briefing. Apixaban for preventing  
691 stroke and systemic embolism in people with nonvalvular atrial fibrillation. Issue date: November  
692 2012. Available from:

- 693 <https://www.nice.org.uk/guidance/ta275/documents/stroke-and-systemic-embolism-prevention->  
694 [nonvalvular-atrial-fibrillation-apixaban-premeeting-briefing2](https://www.nice.org.uk/guidance/ta275/documents/stroke-and-systemic-embolism-prevention-nonvalvular-atrial-fibrillation-apixaban-premeeting-briefing2)
- 695 [10] Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating Randomized  
696 Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT  
697 DUPLICATE Initiative. *Circulation*. 2021 Mar 9.143(10):1002-1013. doi:  
698 10.1161/CIRCULATIONAHA.120.051718. Epub 2020 Dec 17.
- 699 [11] Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, et al. Real world effects of  
700 COPD medications: a cohort study with validation against results from randomised controlled trials.  
701 *Eur Respir J*. 2021 Mar 25;57(3):2001586. doi: 10.1183/13993003.01586-2020.
- 702 [12] Dahabreh IJ, Robins JM, Hernán MA. Benchmarking Observational Methods by Comparing  
703 Randomized Trials and Their Emulations. *Epidemiology*. 2020 Sep;31(5):614-619. doi:  
704 10.1097/EDE.0000000000001231.
- 705 [13] Baptiste PJ, Wong A YS, Schultze A, Clase CM, Leyrat C, Williamson E, et al. Comparative  
706 effectiveness of ARB and ACEi for cardiovascular outcomes and risk of angioedema among different  
707 ethnic groups in England: an analysis in the UK Clinical Practice Research Datalink with emulation of  
708 a reference trial (ONTARGET). medRxiv preprint 2024 Jan 17 doi: doi:10.1101/2024.01.17.24301397.
- 709 [14] Powell EM, Douglas IJ, Gungabissoon U, Smeeth L, Wing K. Real-world effects of medications for  
710 stroke prevention in atrial fibrillation: protocol for a UK population-based non-interventional cohort  
711 study with validation against randomized trial results. *BMJ Open* 2021;11:e042947.  
712 doi:10.1136/bmjopen-2020-04294.
- 713 [15] Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical  
714 Practice Research Datalink (CPRD) Aurum. *Int J Epidemiol*. 2019 Dec 1;48(6):1740-1740g. doi:  
715 10.1093/ije/dyz034.
- 716 [16] Rose M, Beasley N. Center for Drug Evaluation and Research Application number:  
717 202155Orig1s000 MEDICAL REVIEW(S). Available for download at:  
718 [https://www.accessdata.fda.gov/drugsatfda\\_docs/nda/2012/202155Orig1s000TOC.cfm](https://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/202155Orig1s000TOC.cfm)

- 719 [17] Vinereanu D, Stevens SR, Alexander JH, Al-Khatib SM, Avezum A, Bahit MC, et al. Clinical  
720 outcomes in patients with atrial fibrillation according to sex during anticoagulation with apixaban or  
721 warfarin: a secondary analysis of a randomized controlled trial. *Eur Heart J*. 2015 Dec 7;36(46):3268-  
722 75. doi: 10.1093/eurheartj/ehv447.
- 723 [18] Halvorsen S, Atar D, Yang H, De Caterina R, Erol C, et al. Efficacy and safety of apixaban  
724 compared with warfarin according to age for stroke prevention in atrial fibrillation: observations  
725 from the ARISTOTLE trial. *Eur Heart J*. 2014 Jul 21;35(28):1864-72. doi: 10.1093/eurheartj/ehu046.
- 726 [19] Iacus SM, King G, Porro G. Causal Inference without Balance Checking: Coarsened Exact  
727 Matching Political Analysis. 2012;20(1):1-24. doi:10.1093/pan/mpr013
- 728 [20] Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug  
729 effect studies by time-conditional propensity scores. *Pharmacoepidemiol Drug Saf*. 2017  
730 Apr;26(4):459-468. doi: 10.1002/pds.4107.
- 731 [21] Webster-Clark M, Mavros P, Garry EM, Stürmer T, Shmuel S, Young J, et al. Alternative analytic  
732 and matching approaches for the prevalent new-user design: A simulation study.  
733 *Pharmacoepidemiol Drug Saf*. 2022 Jul;31(7):796-803. doi: 10.1002/pds.5446.
- 734 [22] Austin, P.C. (2009), Some methods of propensity-score matching had superior performance to  
735 others: results of an empirical investigation and Monte Carlo simulations. *Biom. J*. 2009 Feb;51(1):  
736 171-84. doi: 10.1002/bimj.200810488
- 737 [23] Woodfield R, Grant I, UK Biobank Stroke Outcomes Group, UK Biobank Follow-Up and Outcomes  
738 Working Group, Sudlow CL. Accuracy of Electronic Health Record Data for Identifying Stroke Cases in  
739 Large-Scale Epidemiological Studies: A Systematic Review from the UK Biobank Stroke Outcomes  
740 Group. *PLoS One*. 2015 Oct 23;10(10):e0140533. doi: 10.1371/journal.pone.0140533.
- 741 [24] Austin P.C, Small DS. The use of bootstrapping when using propensity-score matching without  
742 replacement: A simulation study. *Stat Med*. 2014 Oct 30; 33(24), 4306–19. doi: 10.1002/sim.6276.
- 743 [25] Abadie A., Spiess J. Robust Post-Matching Inference. *Journal of the American Statistical*  
744 *Association*, 117:538, 983-995, doi: 10.1080/01621459.2020.1840383[26] Lobo FS, Wagner S, Gross

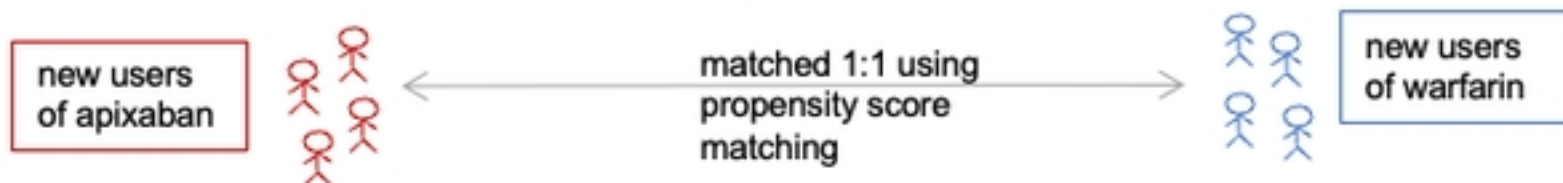
745 CR, Schommer JC. Addressing the issue of channeling bias in observational studies with propensity  
746 scores analysis. *Res SocialAdm Pharm.* 2006 Mar;2(1):143-51. doi:  
747 10.1016/j.sapharm.2005.12.001.[27] Rosenbaum PR, Rubin DB. The central role of the propensity  
748 score in observational studies for causal effects. *Biometrika.* 1983 Apr;70(1):41-55. doi:  
749 10.1093/biomet/70.1.41.  
750 [28] Blake HA, Leyrat C, Mansfield KE, Seaman S, Tomlinson LA, Carpenter J, et al. Propensity scores  
751 using missingness pattern information: a practical guide. *Stat Med.* 2020 May 20;39(11):1641–1657.  
752 doi: 10.1002/sim.8503.  
753 [29] Vinogradova Y, Coupland C, Hill T, Hippisley-Cox J. Risks and benefits of direct oral  
754 anticoagulants versus warfarin in a real world setting: cohort study in primary care. *BMJ.* 2018 Jul 4;  
755 362:k2505 doi:10.1136/bmj.k2505. Erratum in: *BMJ.* 2018 Oct 18;363:k4413.  
756 [30] Larsen TB, Skjøth F, Nielsen PB, Kjældgaard JN, Lip GY. Comparative effectiveness and safety of  
757 non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation:  
758 propensity weighted nationwide cohort study. *BMJ.* 2016 Jun 16;353:i3189. doi:10.1136/bmj.i3189.  
759 [31] Li XS, Deitelzweig S, Keshishian A, Hamilton M, Horblyuk R, Gupta K, et al. Effectiveness and  
760 safety of apixaban versus warfarin in non-valvular atrial fibrillation patients in “real-world” clinical  
761 practice. A propensity-matched analysis of 76,940 patients. *Thromb Haemost.* 2017 Jun  
762 2;117(6):1072-1082. doi:10.1160/TH17-01-0068  
763 [32] Proietti M, Romanazzi I, Romiti GF, Farcomeni A, Lip GYH. Real-World Use of Apixaban for  
764 Stroke Prevention in Atrial Fibrillation: A Systematic Review and Meta-Analysis. *Stroke.* 2018  
765 Jan;49(1):98-106. doi: 10.1161/STROKEAHA.117.018395.  
766 [33] Franklin J, Brigham and Women's Hospital. Replication of the ARISTOTLE Anticoagulant Trial in  
767 Healthcare Claims Data. Available from: <https://www.clinicaltrials.gov/ct2/show/NCT04593030>  
768 [34] Wang SV, Schneeweiss S, RCT-DUPLICATE initiative, Franklin JM, Desai RJ, Feldman W, et al.  
769 Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32  
770 Clinical Trials. *JAMA.* 2023 Apr 25;329(16):1376-1385. doi:10.1001/jama.2023.4221.



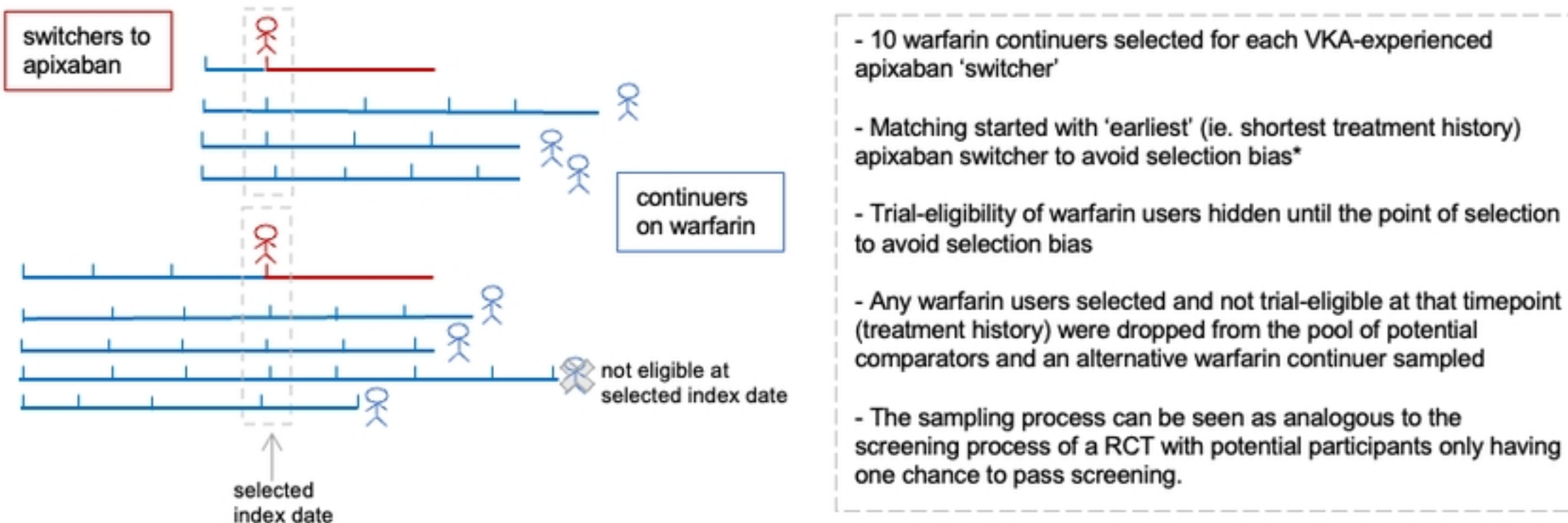
- 771 [35] Pokorney SD, Simon DN, Thomas L, Fonarow GC, Kowey PR, Chang P, et al. Patients' time in  
772 therapeutic range on warfarin among US patients with atrial fibrillation: Results from ORBIT-AF  
773 registry. *Am Heart J*. 2015 Jul;170(1):141-148.e1. doi: 10.1016/j.ahj.2015.03.017.
- 774 [36] Chang SH, Chou IJ, Yeh YH, Chiou MJ, Wen MS, Kuo CT, et al. Association Between Use of Non-  
775 Vitamin K Oral Anticoagulants With and Without Concurrent Medications and Risk of Major Bleeding  
776 in Nonvalvular Atrial Fibrillation. *JAMA*. 2017 Oct 3;318(13):1250-1259. doi:  
777 10.1001/jama.2017.13883.
- 778 [37] Schaefer JK, Li Y, Gu X, et al. Association of Adding Aspirin to Warfarin Therapy Without an  
779 Apparent Indication With Bleeding and Other Adverse Events. *JAMA Intern Med*. 2019;179(4):533–  
780 541. doi:10.1001/jamainternmed.2018.7816
- 781 [38] Almas T, Musheer A, Ejaz A, et al. Efficacy and safety of direct oral anticoagulants with and  
782 without Aspirin: A systematic review and Meta-analysis. *Int J Cardiol Heart Vasc*. 2022 Mar  
783 26;40:101016. doi: 10.1016/j.ijcha.2022.101016.
- 784 [37] Sabir I, Khavandi K, Brownrigg J, Camm AJ. Oral anticoagulants for Asian patients with atrial  
785 fibrillation. *Nat Rev Cardiol*. 2014 May;11(5):290–303 (2014). doi: 10.1038/nrcardio.2014.22.
- 786 [38] Kim HK, Tantry US, Smith SC, et al. The East Asian Paradox: An Updated Position Statement on  
787 the Challenges to the Current Antithrombotic Strategy in Patients with Cardiovascular Disease.  
788 *Thromb Haemost*. 2021 Apr;121(4):422-432. doi: 10.1055/s-0040-1718729.
- 789 [39] Gillott RG, Willan K, Kain K, et al. South Asian ethnicity is associated with a lower prevalence of  
790 atrial fibrillation despite greater prevalence of established risk factors: a population-based study in  
791 Bradford Metropolitan District. *Europace*. 2017 Mar 1;19(3):356-363. doi:  
792 10.1093/europace/euw010.
- 793 [39] Shiekh, S.I., Harley, M., Ghosh, R.E. et al. Completeness, agreement, and representativeness of  
794 ethnicity recording in the United Kingdom's Clinical Practice Research Datalink (CPRD) and linked  
795 Hospital Episode Statistics (HES). *Popul Health Metrics* 21, 3 (2023). [https://doi.org/10.1186/s12963-](https://doi.org/10.1186/s12963-023-00302-0)  
796 023-00302-0

- 797 [40] Paris J Baptiste, Angel YS Wong, Anna Schultze, et al. Comparative effectiveness of ARB and ACEi  
798 for cardiovascular outcomes and risk of angioedema among different ethnic groups in England: an  
799 analysis in the UK Clinical Practice Research Datalink with emulation of a reference trial (ONTARGET).  
800 medRxiv 2024.01.17.24301397; doi: <https://doi.org/10.1101/2024.01.17.24301397>
- 801 [41] Banerjee A, Benedetto V, Gichuru P, Burnell J, Antoniou S, Schilling RJ et al. Adherence and  
802 persistence to direct oral anticoagulants in atrial fibrillation: a population-based study. *Heart*. 2020  
803 Jan;106(2):119-126. doi: [10.1136/heartjnl-2019-315307](https://doi.org/10.1136/heartjnl-2019-315307).

i. **New users of apixaban (VKA-naïve) propensity score matched 1:1 to new users of warfarin.**



ii. **Pool of prevalent warfarin users** containing all warfarin prescription dates in the study period for each warfarin user **sampled without replacement by treatment history** to ensure similar duration of prior VKA exposure for the prevalent users in the apixaban and warfarin CPRD Aurum arms:



iii. **VKA-experienced apixaban users propensity score matched 1:1 to the treatment-history matched comparator pool.** Propensity scores were calculated stratified by treatment history strata and patients matched within treatment history strata.



Fig1

AURUM - Apixaban

AURUM - Warfarin

medRxiv preprint doi: <https://doi.org/10.1101/2024.03.18.24304452>; this version posted March 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

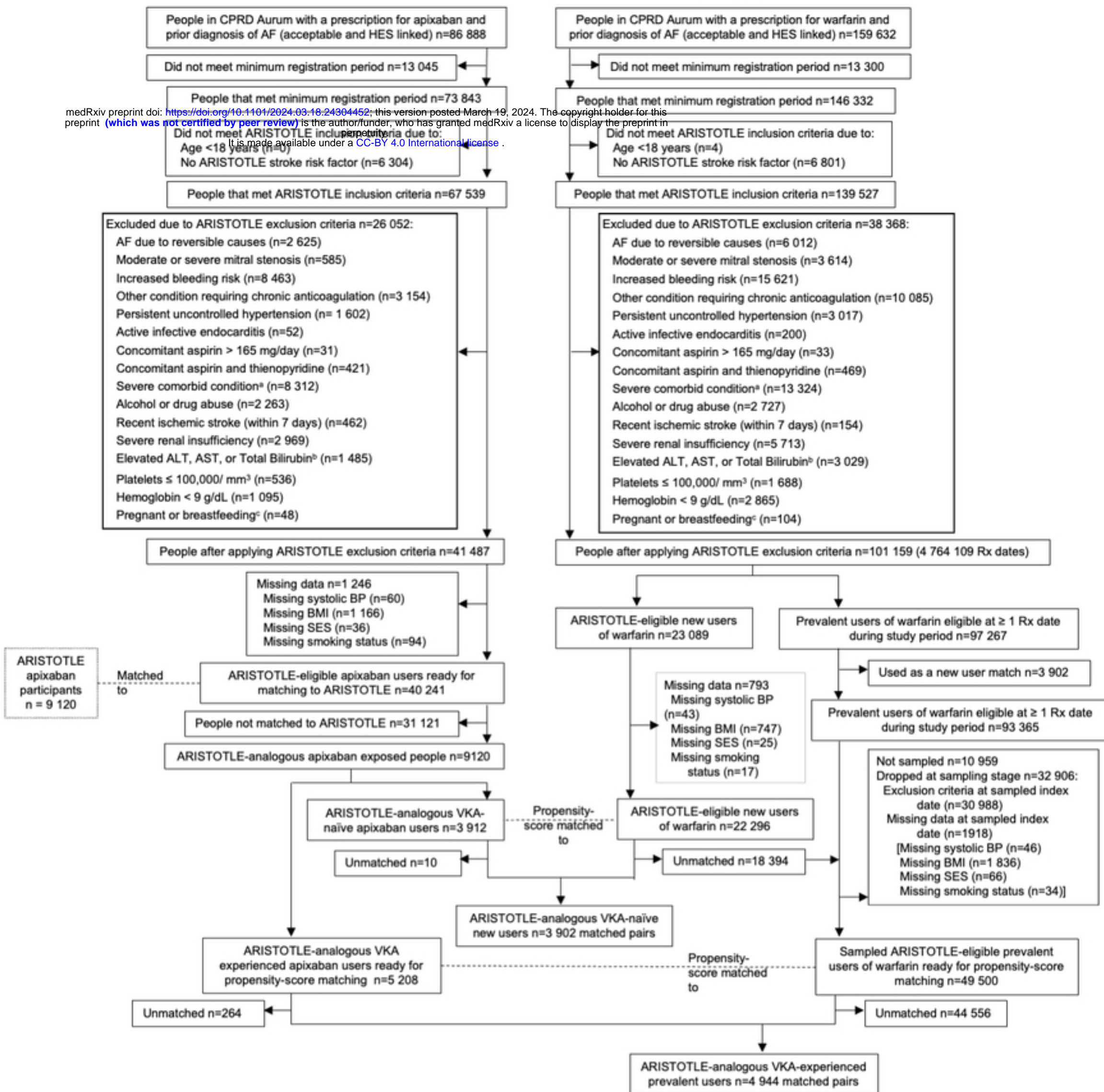


Fig2



Outcome	Rate (%/yr)		HR (95% CI)
	Apx	Warf	
<b>Stroke or systemic embolism</b>			
ARISTOTLE	1.27	1.60	0.79 (0.66,0.95)
CPRD Aurum	1.27	1.29	0.98 (0.82,1.19)
CPRD Aurum TTR <0.75	1.44	1.53	0.94 (0.75,1.19)
CPRD Aurum TTR ≥0.75	1.11	0.75	1.49 (1.13,1.97)
<b>Ischemic or uncertain stroke</b>			
ARISTOTLE	0.97	1.05	0.92 (0.74,1.13)
CPRD Aurum	0.92	0.80	1.13 (0.90,1.41)
CPRD Aurum TTR <0.75	1.02	0.96	1.07 (0.80,1.42)
CPRD Aurum TTR ≥0.75	0.81	0.47	1.74 (1.23,2.45)
<b>Hemorrhagic stroke</b>			
ARISTOTLE	0.24	0.47	0.51 (0.35,0.75)
CPRD Aurum	0.21	0.33	0.67 (0.44,1.01)
CPRD Aurum TTR <0.75	0.23	0.39	0.61 (0.37,1.00)
CPRD Aurum TTR ≥0.75	0.19	0.20	1.04 (0.59,1.85)
<b>Death from any cause</b>			
ARISTOTLE	3.52	3.94	0.89 (0.80,0.998)
CPRD Aurum	4.37	4.20	1.03 (0.93,1.14)
CPRD Aurum TTR <0.75	5.24	5.30	0.99 (0.87,1.12)
CPRD Aurum TTR ≥0.75	3.62	2.06	1.75 (1.49,2.06)

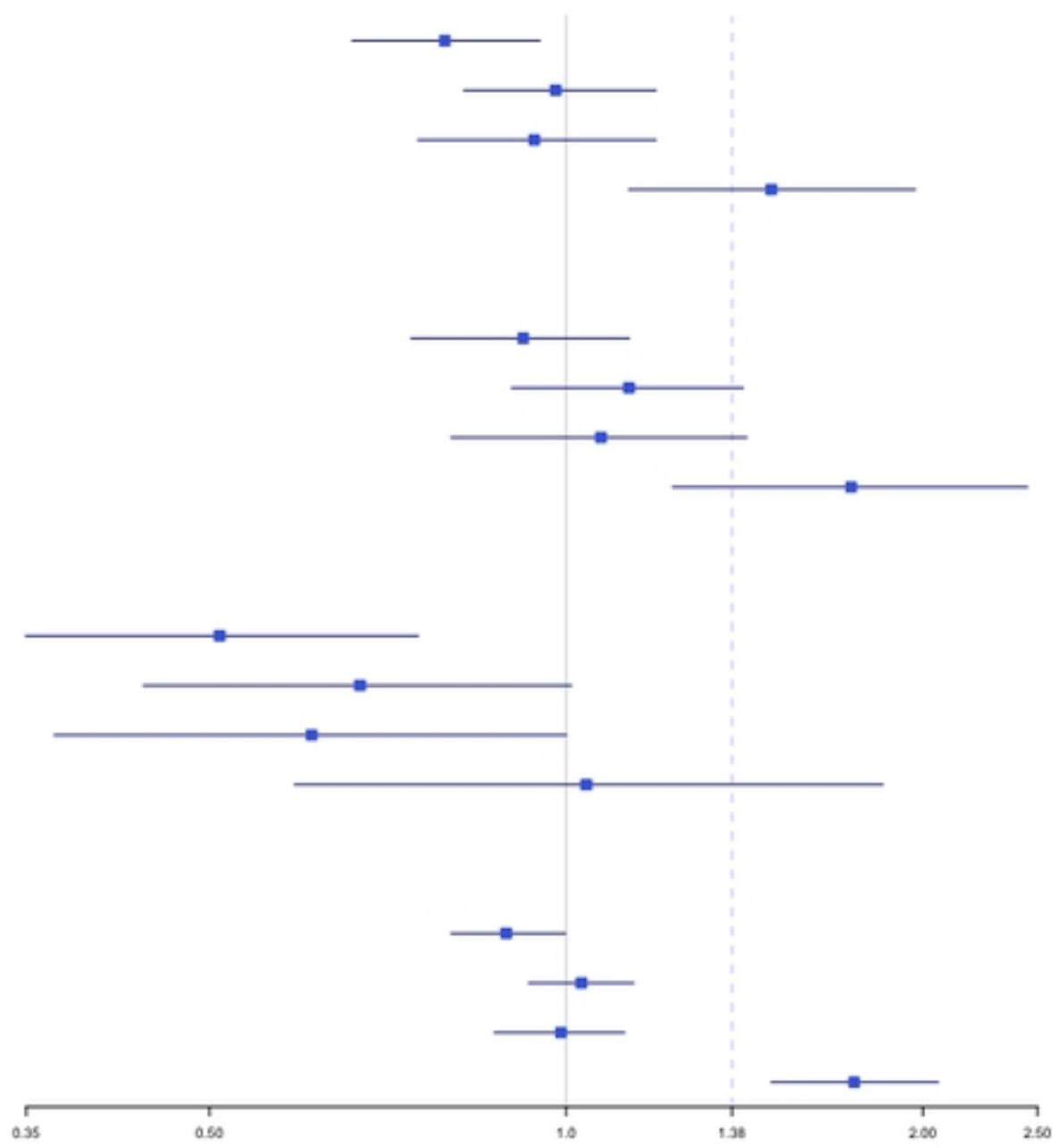


Fig3

Outcome	Rate (%/yr)		HR (95% CI)
	Apx	Warf	
<b>Major bleeding</b>			
ARISTOTLE	2.13	3.09	0.69 (0.60,0.80)
CPRD Aurum	2.45	2.77	0.88 (0.77,1.00)
CPRD Aurum TTR <0.75	2.69	3.85	0.70 (0.59,0.82)
CPRD Aurum TTR ≥0.75	2.22	1.71	1.31 (1.08,1.59)
<b>Intracranial bleeding</b>			
ARISTOTLE	0.33	0.80	0.42 (0.30,0.58)
CPRD Aurum	0.35	0.50	0.71 (0.51,1.00)
CPRD Aurum TTR <0.75	0.36	0.68	0.55 (0.37,0.83)
CPRD Aurum TTR ≥0.75	0.31	0.28	1.13 (0.69,1.83)
<b>Other location bleeding</b>			
ARISTOTLE	1.79	2.27	0.79 (0.68,0.93)
CPRD Aurum	0.60	0.64	0.93 (0.70,1.22)
CPRD Aurum TTR <0.75	0.64	0.91	0.70 (0.50,0.99)
CPRD Aurum TTR ≥0.75	0.56	0.42	1.35 (0.91,2.00)
<b>Gastrointestinal bleeding</b>			
ARISTOTLE	0.76	0.86	0.89 (0.70,1.15)
CPRD Aurum	1.52	1.70	0.88 (0.74,1.04)
CPRD Aurum TTR <0.75	1.71	2.37	0.72 (0.58,0.88)
CPRD Aurum TTR ≥0.75	1.39	1.01	1.35 (1.05,1.73)

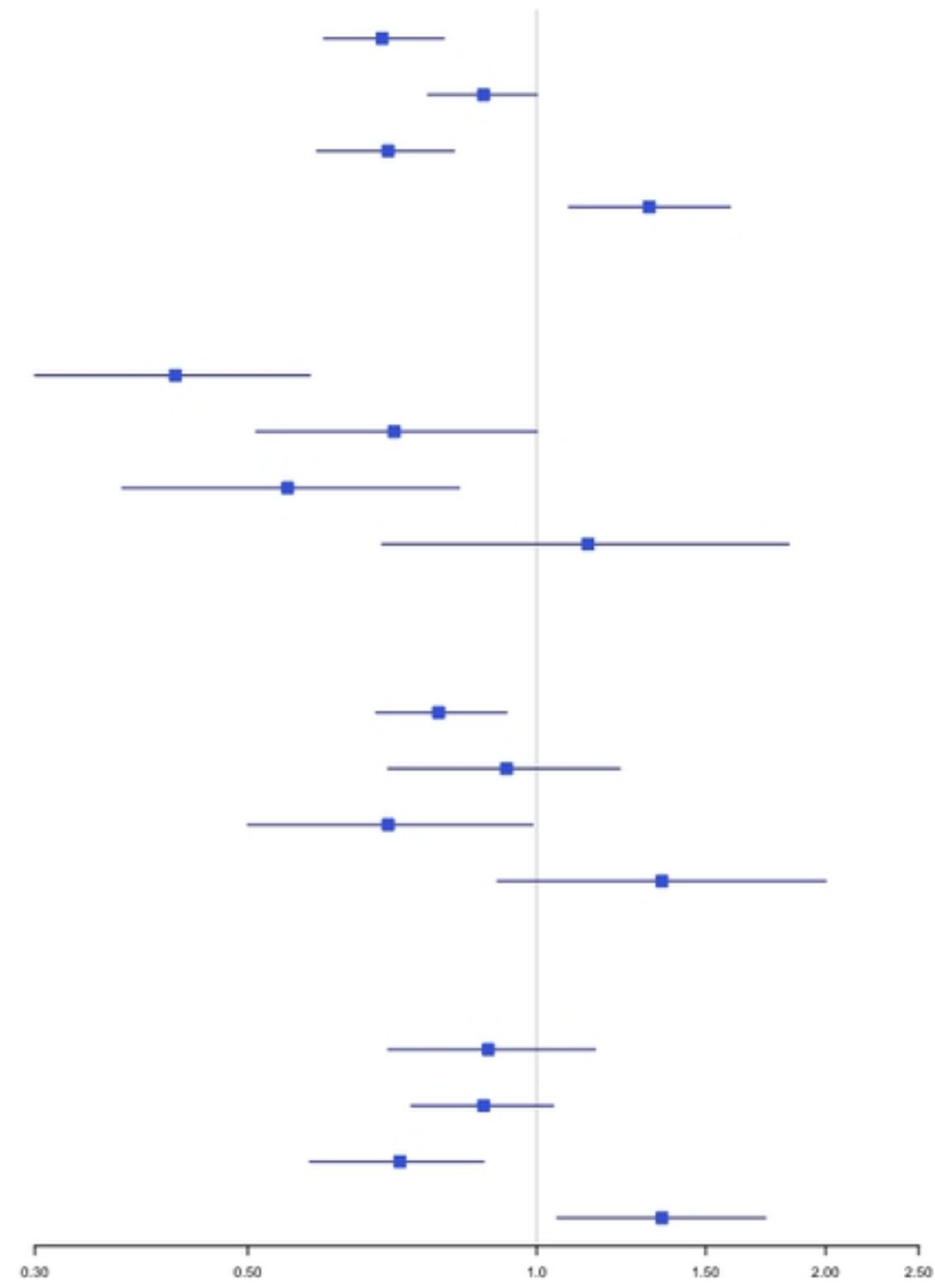


Fig4