openheart

Emulation of ARISTOTLE and ROCKET AF trials in real-world atrial fibrillation patients results in similar efficacy and safety as original landmark trials: insights from the GARFIELD-AF registry

Jelle C L Himmelreich ,[1,2,3] Saverio Virdone,[1] A John Camm,[4] Karen Pieper,[1] Ralf E Harskamp,[2,3] Freek W A Verheugt,[5] Jean-Pierre Bassand,[6] Frank Misselwitz,[7] Antônio C Pereira-Barretto,[8] Frank Cools,[9] Harry Gibbs,[10] Ajay K Kakkar,[1] For the GARFIELD-AF investigators

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/openhrt-2024-002966).

To cite: Himmelreich JCL, Virdone S, Camm AJ, *et al*. Emulation of ARISTOTLE and ROCKET AF trials in real-world atrial fibrillation patients results in similar efficacy and safety as original landmark trials: insights from the GARFIELD-AF registry. *Open Heart* 2025;**12**:e002966. doi:10.1136/openhrt-2024-002966

Received 19 September 2024
Accepted 20 November 2024

For numbered affiliations see end of article.

Correspondence to
Dr Jelle C L Himmelreich; j.c.himmelreich@amsterdamumc.nl

ABSTRACT

Aims This study aimed to determine the robustness, reproducibility and representativeness of the landmark Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (AF) (ARISTOTLE) and Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in AF (ROCKET AF) randomised trials through replication in an observational AF patient registry.

Methods and results Patients from the Global Anticoagulant Registry in the FIELD (GARFIELD)-AF registry treated with apixaban, rivaroxaban or vitamin K antagonist (VKA) were assessed for eligibility for the ARISTOTLE and ROCKET AF trials. HRs of apixaban and rivaroxaban versus comparator for stroke/systemic embolism, major bleeding and all-cause mortality within 2 years follow-up were calculated using propensity score overlap-weighted Cox models. Among GARFIELD-AF patients on apixaban, 2570/3615 (71%) would have been eligible for ARISTOTLE. Among patients using rivaroxaban, 2005/4914 (41%) would have been eligible for ROCKET AF. Eligibility rates were steady over time, with minor differences across medical specialties. Real-world AF patients selected according to trial criteria had lower cardiovascular burden than the original trial participants, especially compared with ROCKET AF. HRs (95% CI) for apixaban versus VKA among ARISTOTLE-eligible users were 0.57 (0.34 to 0.94) for stroke/systemic embolism, 0.76 (0.48 to 1.20) for major bleeding and 0.89 (0.70 to 1.12) for all-cause mortality. Among ROCKET AF-eligible rivaroxaban users, HRs for rivaroxaban versus VKA were 0.90 (0.57 to 1.43), 0.92 (0.59 to 1.43) and 0.86 (0.69 to 1.08), respectively. All safety and efficacy estimates were similar to those in the original trials.

Conclusion Real-world representativeness of the selection criteria was greater for ARISTOTLE than ROCKET AF. The pivotal randomised trials of apixaban and rivaroxaban versus warfarin can be successfully emulated in real-world AF patients by applying trial-

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The landmark trials of apixaban and rivaroxaban—Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation (AF) (ARISTOTLE) and Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial (ROCKET) AF—demonstrated the efficacy and safety of each compared with vitamin K antagonist treatment. However, it is unknown if the results of these trials can be replicated in the real-world AF patient population.

WHAT THIS STUDY ADDS

⇒ Using data from the international Global Anticoagulant Registry in the FIELD-AF registry, this study shows that nearly three-quarters of real-world AF patients treated with apixaban would have been eligible for ARISTOTLE, while only 40% would have been eligible for ROCKET AF. The landmark trial results for stroke/systemic embolism, all-cause mortality and major bleeding in apixaban-treated or rivaroxaban-treated patients compared with vitamin K antagonist-treated patients were successfully emulated in real-world patients.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Randomised controlled trials remain the standard for assessing anticoagulant treatments. However, real-world observational data can be used to emulate trial outcomes when the data are high quality and appropriate non-random treatment allocation methodology is used.

specific selection criteria and appropriate methodology for non-randomised treatment allocation.
Trial registration number NCT01090362.

INTRODUCTION

Atrial fibrillation (AF) is associated with an increased risk of stroke.[1] Effective oral anticoagulation (OAC) treatment, traditionally by vitamin K antagonists (VKA), decreases the burden of stroke and mortality in at-risk AF patients.[2] Clinical research over the past decade has shown that non-vitamin K OACs (NOACs) are not only non-inferior to VKAs but offer a superior balance between reducing risk of stroke and increasing risk of bleeding.[3–6] NOACs have thus become widely recommended for stroke prophylaxis in patients with AF[2 7 8] and are currently the most common anticoagulant therapy.[9]

The external validity of key NOAC randomised controlled trials (RCTs) relies, in part, on the generalisability of their inclusion and exclusion criteria to real-world settings.[10] Two factor Xa inhibitor trials, ARISTOTLE (Apixaban for Reduction in Stroke and Other Thromboembolic Events in AF) and ROCKET AF (Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in AF) used considerably different inclusion criteria, resulting in trial populations with markedly different baseline characteristics.[3 4] This has sparked debate on how relevant these studies were for real-world AF populations.[11 12]

Postmarketing surveillance studies and registry data have confirmed the overall safety and efficacy of NOACs for stroke prevention in real-world AF patients.[13 14] However, the landmark randomised ARISTOTLE and ROCKET AF trials have not been replicated or emulated in large, prospectively collected observational datasets using the original trial inclusion and exclusion criteria. The Global Anticoagulant Registry in the FIELD AF (GARFIELD-AF) is the world's largest registry of treatment and outcomes in newly diagnosed non-valvular AF patients.[15] Patients were recruited during the period when NOACs entered clinical practice. This provided a unique opportunity to emulate NOAC versus VKA trials and assess the generalisability of the ARISTOTLE and ROCKET AF selection criteria to patients from a real-world AF registry.

METHODS

The GARFIELD-AF registry

GARFIELD-AF is a prospective, observational study of patients with non-valvular AF from 1215 sites across 35 countries across the globe.[15] Briefly, adults diagnosed with AF within the preceding 6 weeks, excluding cases with valvular or transient reversible AF, and with at least one additional risk factor for stroke as assessed by local practitioners, were eligible for enrolment. Patients were enrolled consecutively to minimise recruitment bias, in five separate sequential cohorts between March 2010 and August 2016. The choice of treatment was at the discretion of local practitioners. All participants were followed up to 2 years after study enrolment and the database has been closed.

Patient selection

Analysis involved patients from GARFIELD-AF cohorts 3–5, recruited during April 2013–August 2016 (n=34 903), whose baseline OAC treatment included either apixaban, rivaroxaban or VKA. Cohorts 1 and 2 were not included because individual drug names were not recorded for patients enrolled and, during their time of enrolment (ie, 2010–2013), NOACs had not yet been introduced in many participating countries. Baseline treatment was defined as a participant's first registered stroke prophylaxis, regardless of OAC dosage or concomitant antiplatelet (AP) treatment. Patients with incomplete data on baseline treatment or clinical outcomes during follow-up were excluded.

Baseline data collection

Oversight of operations and data management of GARFIELD-AF were performed by the Thrombosis Research Institute (TRI; London, UK). A 20% portion of all eCRFs were monitored against source documentation, there was an electronic audit trail for all data modifications, and critical variables were subjected to additional audit.[15 16] Patient baseline data were assessed at enrolment into GARFIELD-AF. Data for components of the $CHADS_2$, CHA_2DS_2-VASc, HAS-BLED and GARFIELD-AF risk stratification schemes were collected and calculated retrospectively.[17–19] Fluctuations in the international normalised ratio were excluded from HAS-BLED score calculations since patients had not previously received AF thromboprophylaxis. Follow-up data were collected at 4-month intervals up to 24 months. Data for this report were extracted from the study database on 30 June 2019.

Outcomes

The primary efficacy outcome was the composite of stroke (haemorrhagic, ischaemic or unknown type of stroke) or systemic embolism (SE); the secondary efficacy outcome was all-cause mortality. The primary safety outcome was major bleeding, defined according to International Society on Thrombosis and Haemostasis (ISTH) criteria[20] as clinically overt bleeding associated with fall in haemoglobin of ≥20 g/L, or associated with transfusion of ≥2 units of packed red blood cells or whole blood or bleeding in a critical site, namely intracranial (spontaneous intracerebral, intraventricular, subarachnoidal, subdural, epidural), intraspinal, pericardial, intra-articular, intramuscular with compartment syndrome, or retroperitoneal, or leading to a fatal outcome. Clinical events were defined prior to patient enrolment as reported previously.[15]

Landmark trial eligibility

Review of the trial protocols for ARISTOTLE[3] and ROCKET AF[4] resulted in 8 distinct inclusion and 49 exclusion criteria across both (online supplemental table S1). These criteria were matched with GARFIELD-AF eCRFs, or operationalised where possible as described in the table footnotes. GARFIELD-AF patients treated with the respective NOAC or dose-adjusted warfarin were characterised as eligible for ARISTOTLE or ROCKET AF if one or more trial inclusion criteria and no exclusion criteria were considered to be present at baseline.

The ARISTOTLE and ROCKET AF protocols also contained specific inclusion targets: ARISTOTLE aimed for <40% of patients with prior VKA use, whereas ROCKET AF limited the percentage of patients with the lowest stroke risk (defined as no history of stroke, transient ischaemic attack (TIA) or SE and two risk factors for stroke, equivalent to a $CHADS_2$ score ≥2 to 10%.[21] These percentages were assessed in the ARISTOTLE- and ROCKET AF-eligible GARFIELD-AF populations, though the eligibility criteria were not modified to better reflect the aimed targets of the two trials.

Statistical analysis

Numbers and percentages are reported for categorical variables; medians and IQR are reported for continuous variables. Missing data percentages are provided for baseline variables (online supplemental table S2). Crude event rates were estimated using a Poisson model-adjusted analyses of clinical outcomes incorporated imputed missing values from patients' baseline characteristics with the multivariable imputation by chained equations method.[22] HRs estimates were obtained combining estimates across five imputed datasets.

To assess generalisability of patient selection to real-world AF populations, annual trial eligibility rates were calculated, both for the overall sample and by care setting specialty. The proportion of GARFIED-AF patients which would have attained the required inclusion criteria for ARISTOTLE and ROCKET AF are reported. Finally, key baseline characteristics of the ARISTOTLE and ROCKET AF participants and of trial-eligible patients from the GARFIELD-AF registry are provided.

In a real-world setting, patient baseline characteristics can affect treatment allocation. Propensity weighting is an important technique to emulate the conditions of a randomised clinical trial.[23] Propensity scores of apixaban versus VKA and of rivaroxaban versus VKA were generated using logistic regressions. Treatment comparisons were conducted using Cox proportional-hazards models, with the propensity method of overlap weighting applied to balance covariates within the population.[24] This method outperforms inverse probability of treatment weighting when the comparator groups are initially very different. In contrast to propensity score matching, propensity score weighting has no impact on sample size. Variables included in the propensity score weighting scheme are reported in online supplemental table S3. In order to assess the balance of confounders between VKA and NOAC in the trial emulation analyses, the absolute standardised differences of variables before and after propensity score weighting were graphically assessed. Time-to-event analyses included patients from time of enrolment until the first of the outcome of interest, loss to follow-up, early withdrawal or 2 years of follow-up. Data analysis was carried out at the TRI using SAS Enterprise Guide V.8.2 (SAS Institute).

RESULTS

Trial eligibility

All inclusion criteria and a number of key exclusion criteria of both trials were matched (online supplemental table S1). Among the trial exclusion criteria that could not be verified from GARFIELD-AF data, a considerable number were either already excluded from GARFIELD-AF (eg, reversible cause for AF) or clinically unlikely to have been incorporated in a database of real-world newly diagnosed AF patients treated with apixaban or rivaroxaban (eg, planned or recent major surgery, or active or recent major bleeding). Within GARFIELD-AF, application of the inclusion and exclusion criteria from ARISTOTLE and ROCKET AF resulted in 2570 (71%) ARISTOTLE-eligible apixaban users and 2005 (41%) ROCKET AF-eligible rivaroxaban users. Among VKA-treated patients, 8005 (68%) were eligible for ARISTOTLE and 4368 (37%) were eligible for ROCKET AF (figure 1). All but two ROCKET AF-eligible VKA users (0.05%) were also included in the ARISTOTLE group. Eligibility rates among NOAC users for their landmark trial were steady over time, with only minor differences across specialties (online supplemental figure S1).

The rate of prior VKA use in ARISTOTLE-eligible VKA and apixaban users was 9%, within the <40% target set by the ARISTOTLE protocol. However, the proportion of ROCKET AF-eligible patients with low stroke risk ($CHADS_2$=2) was over 60% in both the rivaroxaban and VKA arm of GARFIELD-AF, substantially exceeding the ROCKET AF target maximum 10%. Consequently, the $CHADS_2$ mean score of rivaroxaban-treated patients in GARFIELD-AF was lower than in the original trial (figure 2).

The reasons for trial ineligibility among GARFIELD-AF patients are shown in online supplemental table S4. Of the GARFIELD-AF patients ineligible for ROCKET AF or ARISTOTLE, the most common reasons for exclusion were no prior stroke or TIA, or for having less than two or no risk factors for stroke at baseline. Additionally, approximately one-third of patients ineligible for ARISTOTLE had a high risk of bleeding at baseline.

Patient characteristics

Baseline data of trial-eligible GARFIELD-AF patients are shown in table 1. Median $CHADS_2$, CHA_2DS_2-VASc and HAS-BLED risk scores among GARFIELD-AF participants were similar in all groups, regardless of

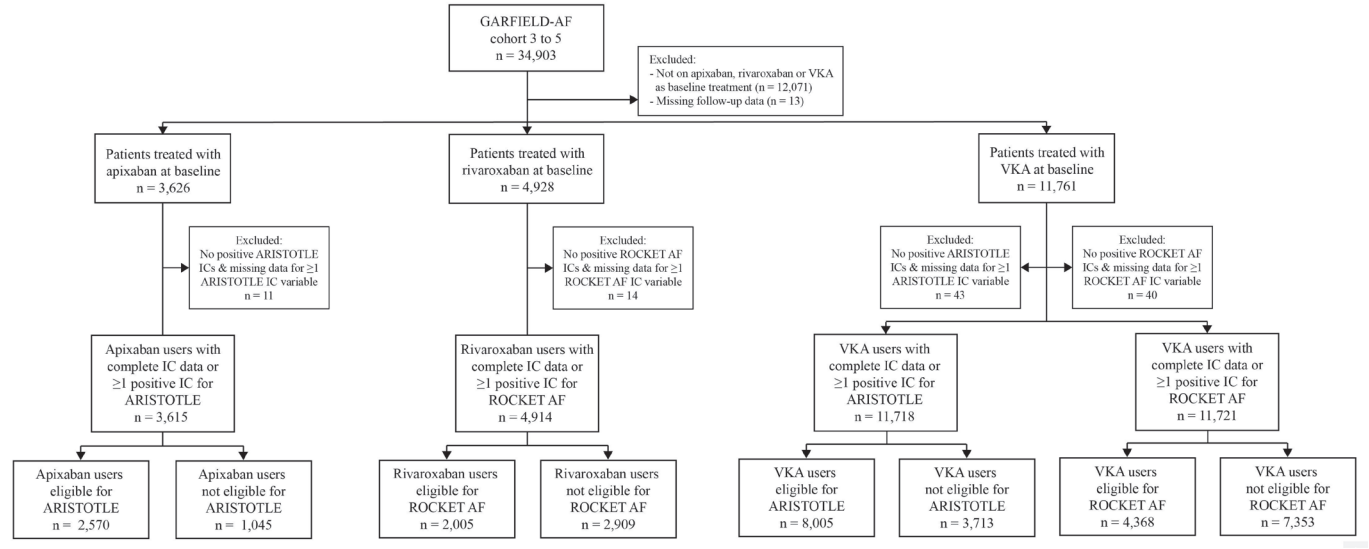

Figure 1 Flow chart for the selection of study population. AF, atrial fibrillation; ARISTOTLE, Apixaban for Reduction in Stroke and Other Thromboembolic Events; GARFIELD, Global Anticoagulant Registry in the FIELD; ROCKET, Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial; VKA, vitamin K antagonist.

treatment. ROCKET AF-eligible GARFIELD-AF patients had higher median 2-year expected risk of death, stroke and bleeding compared with ARISTOTLE-eligible patients, independently of the OAC they were prescribed at baseline. Compared with NOAC users, VKA users in GARFIELD-AF suffered more frequently from diabetes, heart failure, acute coronary syndrome, vascular and carotid occlusive disease and were more often receivers of concomitant AP treatment. Conversely, moderate to severe CKD and carotid occlusive disease were more

	ARISTOTLE			ROCKET AF		
	GAF, apixaban	GAF, VKA	Original trial[1]	GAF, rivaroxaban	GAF, VKA	Original trial[2]
Age (median, IQR)	76 (69-81)	71 (64-77)	70 (63-76)	76 (70-81)	75 (66-79)	73 (65-78)
Female (%)	48	47	35	47	47	40
Prior stroke/TIA/SE (%)	14	13	19	16	18	55
Congestive heart failure (%)[3]	22	25	35	29	30	62
Hypertension (%)[4]	66	74	87	97	98	91
Diabetes mellitus (%)[5]	24	28	25	36	38	40
$CHADS_2$ (mean ± SD)	2.0 ± 1.0	1.9 ± 0.9	2.1 ± 1.1	2.6 ± 0.8	2.5 ± 0.8	3.5 ± 1.0
$CHADS_2$ distribution[6]						

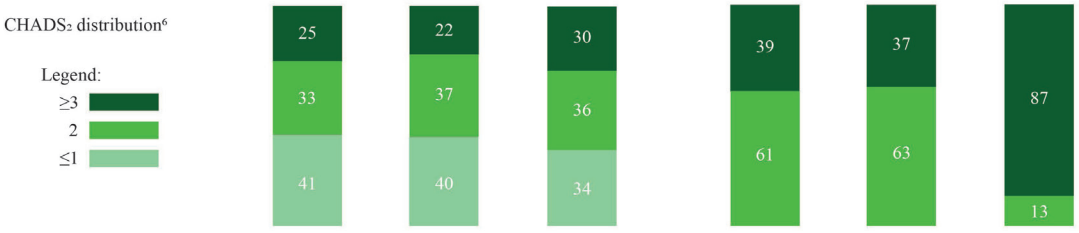

Figure 2 Comparison of key clinical characteristics of trial-eligible patients in GARFIELD-AF and participants in the original trials. Since no relevant differences in characteristics were seen between study arms in either ARISTOTLE or ROCKET AF trial, data from treatment groups in the original trials are presented as the mean. Definitions of some characteristics differed slightly between the trials as explained in the footnotes. $CHADS_2$ scores were applied as in the respective original ARISTOTLE and ROCKET AF trials. [1]Granger et al., 2011. [2]Patel et al., 2011. [3]Defined in ARISTOTLE as a history of heart failure or left ventricular ejection fraction ≤40% and in ROCKET AF as a history of heart failure or left ventricular ejection fraction ≤35%. [4]Defined in ARISTOTLE as hypertension requiring pharmacological treatment and in ROCKET AF as a history of hypertension or hypertension treatment. [5]Defined in ARISTOTLE as a history of diabetes mellitus, and in ROCKET AF as a history of diabetes mellitus or use of antidiabetic medication. [6]Using the $CHADS_2$ element definitions as applied in the respective original ARISTOTLE and ROCKET AF trials. AF, atrial fibrillation; ARISTOTLE, Apixaban for Reduction in Stroke and Other Thromboembolic Events in AF; $CHADS_2$, Congestive heart failure, Hypertension, Age≥75 years, Diabetes, previous Stroke score; GARFIELD, Global Anticoagulant Registry in the FIELD; ROCKET AF, Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in AF; SE, systemic embolism; TIA, transient ischaemic attack; VKA, vitamin K antagonist.

Table 1 Baseline characteristics by considered trial and OAC treatment at baseline in GARFIELD-AF patients eligible for the respective trials

Baseline characteristics	ARISTOTLE		ROCKET AF	
	Apixaban (n=2570)	VKA (n=8005)	Rivaroxaban (n=2005)	VKA (n=4368)
Female sex, n (%)	1240 (48.2)	3747 (46.8)	949 (47.3)	2072 (47.4)
Age, median (Q1; Q3), years	76.0 (69.0; 81.0)	71.0 (64.0; 77.0)	76.0 (70.0; 81.0)	75.0 (66.0; 79.0)
Ethnicity, n (%)				
Caucasian	1513 (60.8)	5296 (67.3)	1389 (69.3)	2930 (68.3)
Hispanic/Latino	70 (2.8)	666 (8.5)	137 (6.8)	405 (9.4)
Asian	859 (34.5)	1730 (22.0)	413 (20.6)	862 (20.1)
Afro-Caribbean/mixed/other	46 (1.8)	174 (2.2)	31 (1.5)	94 (2.2)
Body mass index, median (Q1; Q3), kg/m²	26.2 (23.4; 29.8)	27.8 (24.6; 32.0)	27.5 (24.3; 31.2)	27.6 (24.5; 31.7)
Systolic blood pressure, median (Q1; Q3), mm Hg	130.0 (120.0; 142.0)	130.0 (120.0; 144.0)	132.0 (120.0; 142.0)	130.0 (120.0; 142.0)
Diastolic blood pressure, median (Q1; Q3), mm Hg	79.0 (70.0; 85.0)	80.0 (70.0; 88.0)	80.0 (70.0; 84.5)	80.0 (70.0; 85.0)
Pulse, median (Q1; Q3), bpm	84.0 (70.0; 107.0)	84.0 (72.0; 100.0)	82.0 (70.0; 100.0)	83.0 (71.0; 100.0)
Type of atrial fibrillation, n (%)				
Permanent	340 (13.2)	1398 (17.5)	327 (16.3)	878 (20.1)
Persistent	392 (15.3)	1358 (17.0)	294 (14.7)	669 (15.3)
Paroxysmal	895 (34.8)	1736 (21.7)	602 (30.0)	891 (20.4)
New onset (unclassified)	943 (36.7)	3513 (43.9)	782 (39.0)	1930 (44.2)
Care specialty, n (%)				
Internal medicine/neurology/geriatrics	486 (18.9)	1558 (19.5)	438 (21.8)	895 (20.5)
Cardiology	1826 (71.1)	5122 (64.0)	1331 (66.4)	2694 (61.7)
Primary care/general practice	258 (10.0)	1325 (16.6)	236 (11.8)	779 (17.8)
Care setting, n (%)				
Hospital	1190 (46.3)	4643 (58.0)	918 (45.8)	2518 (57.6)
Office/anticoagulation clinic/thrombosis centre	1142 (44.4)	2514 (31.4)	926 (46.2)	1419 (32.5)
Emergency room	238 (9.3)	848 (10.6)	161 (8.0)	431 (9.9)
Medical history, n (%)				
Heart failure	567 (22.1)	2046 (25.6)	594 (29.6)	1353 (31.0)
Acute coronary syndromes	227 (8.9)	798 (10.0)	211 (10.6)	486 (11.2)
Vascular disease*	491 (19.3)	1947 (24.6)	512 (25.8)	1153 (26.7)
Carotid occlusive disease	90 (3.6)	176 (2.2)	60 (3.1)	110 (2.6)
VTE	53 (2.1)	260 (3.3)	54 (2.7)	117 (2.7)
Prior stroke/TIA/SE	366 (14.4)	1032 (13.0)	324 (16.3)	758 (17.5)
Hypertension	2082 (81.2)	6761 (84.7)	1715 (85.6)	3698 (84.9)
Hypercholesterolaemia	1114 (44.3)	3506 (46.0)	955 (49.4)	1993 (47.8)
Diabetes	621 (24.2)	2219 (27.7)	712 (35.5)	1629 (37.3)
Moderate to severe CKD†	219 (8.8)	469 (6.3)	201 (10.4)	293 (7.2)
Current smoker, n (%)	181 (7.9)	694 (9.4)	154 (7.7)	360 (9.0)
AP treatment, n (%)	372 (14.7)	1665 (20.8)	328 (16.7)	994 (22.8)
$CHADS_2$ score‡, median (Q1; Q3)	2.0 (1.0; 3.0)	2.0 (1.0; 2.0)	2.0 (2.0;3.0)	2.0 (2.0; 3.0)
CHA_2DS_2-VASc score, median (Q1; Q3)	4.0 (3.0; 4.0)	3.0 (3.0; 4.0)	4.0 (3.0; 5.0)	4.0 (3.0; 5.0)
HAS-BLED score§, median (Q1; Q3)	1.0 (1.0; 2.0)	1.0 (1.0; 2.0)	1.0 (1.0; 2.0)	1.0 (1.0; 2.0)
GARFIELD-AF stroke score, median (Q1; Q3)¶	1.4 (1.0; 1.9)	1.5 (1.1; 2.0)	1.5 (1.2; 1.9)	1.7 (1.3; 2.1)
GARFIELD-AF bleeding score, median (Q1; Q3)¶	1.7 (1.2; 2.2)	2.1 (1.5; 2.7)	1.8 (1.4; 2.3)	2.3 (1.7; 2.9)
GARFIELD-AF death score, median (Q1; Q3)¶	4.6 (2.9; 7.7)	5.0 (3.0; 7.7)	5.6 (3.6; 8.5)	6.0 (4.0; 8.8)

Continued

Table 1 Continued

Baseline characteristics	ARISTOTLE		ROCKET AF	
	Apixaban (n=2570)	VKA (n=8005)	Rivaroxaban (n=2005)	VKA (n=4368)

*Defined as peripheral artery disease and/or coronary artery disease.
†CKD was classified according to National Kidney Foundation guidelines: moderate-to-severe (stages 3–5), mild (stages 1 and 2) or none.
‡Using the $CHADS_2$ element definitions as applied in the respective original ARISTOTLE and ROCKET AF trials.
§The risk factor 'Labile INRs' is not included in the HAS-BLED score as it is not collected at baseline. As a result, the maximum HAS-BLED score at baseline is 8 points (not 9). Ethnicity was classified by the investigator in agreement with the patient.
¶Expected probability of non-haemorrhagic stroke/SE, major bleeding or death within 2 years follow-up.[19]
AF, atrial fibrillation; AP, antiplatelet; ARISTOTLE, Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation; CKD, chronic kidney disease; GARFIELD, Global Anticoagulant Registry in the FIELD; OAC, oral anticoagulation; ROCKET, Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial; SE, systemic embolism; TIA, transient ischaemic attack; VKA, vitamin K antagonist; VTE, venous thromboembolism.

prevalent in NOAC-treated compared with VKA-treated patients eligible for the same trial.

Among ARISTOTLE-eligible GARFIELD-AF patients, the median age of apixaban users was greater than that of VKA users (76 vs 71 years, respectively). Apixaban-treated patients were also more often of Asian ethnicity (34.5% vs 22.0%) and suffered more frequently from carotid occlusive disease (3.6% vs 2.2%) or moderate to severe CKD (8.8% vs 6.3%), and less frequently from other comorbidities such as heart failure (22.1% vs 25.6%) and vascular disease (19.3% vs 24.6%).

The average age of ROCKET AF-eligible patients in GARFIELD-AF was similar in both treatment arms (rivaroxaban: 76 years, VKA: 75 years), as was the proportion of females (both 47%). However, rivaroxaban users had a higher prevalence of moderate to severe CKD (10.4% vs 7.2%) and a higher proportion of paroxysmal AF patients (30.0% vs 20.4%) compared with VKA users. ARISTOTLE-eligible GARFIELD-AF patients using apixaban and VKA users were older (76 and 71 years, respectively) than patients in the actual ARISTOTLE trial (70 years), and more often female (48% and 47% vs 35%). Key clinical characteristics of GARFIELD-AF and original trial participants are shown in figure 2. ARISTOTLE-eligible patients in GARFIELD-AF had lower prevalence of prior stroke/TIA/SE, heart failure and hypertension, and thus, a lower mean $CHADS_2$ score (apixaban: 2.0±1.0, VKA: 1.9±0.9) compared with ARISTOTLE participants (2.1±1.1).

The median age of ROCKET AF-eligible GARFIELD-AF patients in both treatment arms (rivaroxaban: 76 years, VKA: 75 years) was higher than those in the ROCKET AF trial (73 years), and so was the proportion of females (47% vs 40%, respectively). Compared with the original trial, ROCKET AF-eligible GARFIELD-AF patients had a slightly lower prevalence of diabetes (36–38% vs 40%) and higher rate of hypertension (97%–98% vs 91%), but a two times lower prevalence of heart failure, and three times lower prevalence of prior stroke/TIA/SE. Consequently, there was a stark difference in $CHADS_2$ distribution, with under 40% with a $CHADS_2$ score≥3 in ROCKET AF-eligible GARFIELD-AF patients compared with 87% in the original trial.

Trial replication in a real-world setting

Emulating ARISTOTLE in GARFIELD-AF resulted in an adjusted HR (95% CI) of 0.57 (0.34 to 0.94) for stroke/SE, 0.76 (0.48 to 1.20) for major bleeding and 0.89 (0.70 to 1.12) for all-cause mortality (figure 3). In the emulated ROCKET AF trial, adjusted HRs for stroke/SE, major bleeding and all-cause mortality were 0.90 (0.57 to 1.43), 0.92 (0.59 to 1.43) and 0.86 (0.69 to 1.08), respectively. The adjusted HRs in observational data showed considerable overlap with results of the original trials in all selected outcomes.

Missingness proportion was low (<3%) for most variables included in the propensity score weighting scheme, with the exception of vital signs (5%–8%), lifestyle factors (10%–15%) and body mass index (~20%) (online supplemental table S2). Online supplemental figures S2 and S3 show that all confounders used in the propensity score were balanced between VKA and NOAC after weighting, indicating that the emulated 'treatment arms' were comparable for all confounders after applying the weighting scheme. Crude event rates and unadjusted HRs for GARFIELD-AF patients eligible for ARISTOTLE and ROCKET AF are shown in online supplemental table S5.

DISCUSSION

Our study shows that the ARISTOTLE and ROCKET AF selection criteria provide limited representativeness for real-world patients with newly diagnosed non-valvular AF. The relative effectiveness of NOAC versus VKA was comparable with the outcomes of the original trials when emulating ARISTOTLE and ROCKET AF and applying the original trial's selection criteria in the GARFIELD-AF registry.

The observation that selection criteria used in RCTs limit the generalisability of their results to real-world target populations has been made before.[25–27] Specifically, eligibility rates in the AF patient registries for ARISTOTLE

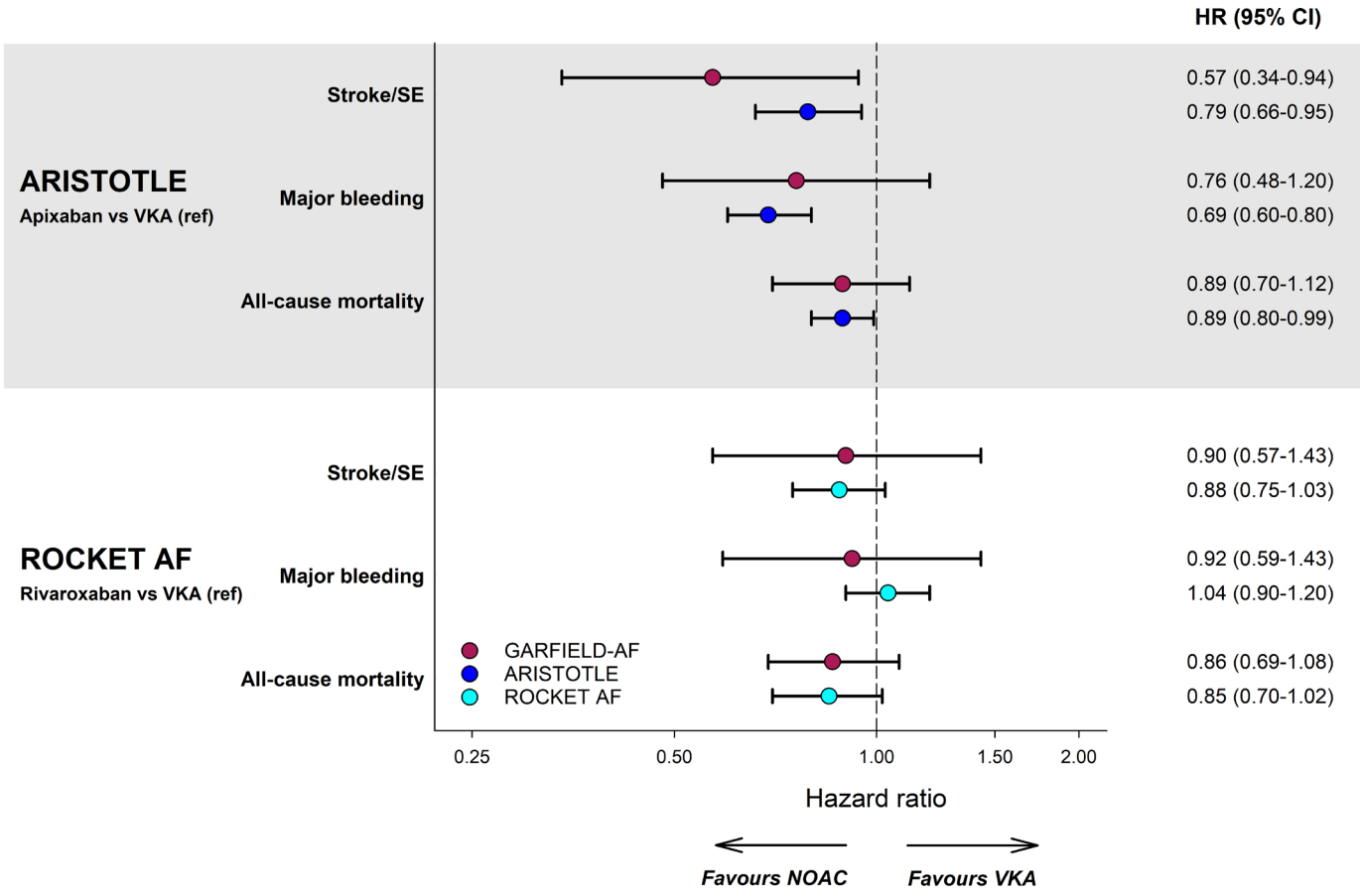

Figure 3 Adjusted HRs comparing NOAC versus VKA. Shown are selected outcomes at 2 years of follow-up in apixaban users eligible for ARISTOTLE (top) and rivaroxaban users eligible for ROCKET AF (bottom), using the VKA users in each group as reference. HRs were obtained using an overlap-weighted Cox model. Variables included in the weighting scheme were country and cohort enrolment, sex, age, ethnicity, type of AF, care setting specialty and location, congestive heart failure, acute coronary syndromes, vascular disease, carotid occlusive disease, prior stroke/transient ischaemic attack/SE, prior bleeding, VTE, hypertension, hypercholesterolaemia, diabetes, moderate to severe chronic kidney disease, hyperthyroidism, hypothyroidism, current smoking, heavy alcohol consumption, body mass index, heart rate, systolic and diastolic blood pressure at diagnosis, and baseline antiplatelet use. AF, atrial fibrillation; ARISTOTLE, Apixaban for Reduction in Stroke and Other Thromboembolic Events; GARFIELD, Global Anticoagulant Registry in the FIELD; NOAC, non-vitamin K antagonist; ROCKET, Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial; SE, systemic embolism; VKA, vitamin K antagonist.

(42%–71%) and ROCKET AF (36%–58%).[11 28–32] This may be due to the specific exclusion and inclusion targets applied in ARISTOTLE and ROCKET AF, resulting in patient groups with higher cardiovascular burden than the average real-world AF patient, especially for ROCKET AF. We found similar rates in GARFIELD-AF: 71% eligibility for ARISTOTLE vs 41% for ROCKET-AF. What sets our study apart is the use of a worldwide registry of patients recruited from sites representative for each country, rather than being restricted to a specific medical specialty, care setting or country.

In addition, we provide an emulation of the original NOAC versus VKA trials that faithfully reproduces the trial's eligibility criteria in the same dataset and produces similar estimates for safety and efficacy. This, rather than observing safety and efficacy in real-world (retrospective) data of otherwise unselected apparent NOAC users, was possible because of the unique size and quality of the GARFIELD-AF data. Three observational studies evaluated the effectiveness and safety in real-world patients but were limited to rivaroxaban and follow-up periods shorter than in ROCKET AF.[33–35] Using discrete element simulation to match trial characteristics, Zhang *et al* showed consistently lower rates of stroke/SE and major bleeding in the rivaroxaban arms of the observational studies compared with ROCKET AF.[36] Moreover, rivaroxaban was non-inferior or even superior to VKA in all four studies. Hence, using a different method and registry data, our new findings confirm these previous similar results while also emulating the ARISTOTLE trial.

In a recent analysis of GARFIELD-AF, we found differences in the risk of stroke/SE when the selection criteria of ARISTOTLE were applied to apixaban-treated or rivaroxaban-treated patients in the registry, suggesting that selection criteria can impact trial outcomes.[37] Our current study complements this by showing that while the

criteria of ARISTOTLE and ROCKET AF have different levels of relevance to the real-world AF population, their outcomes can be emulated when their selection criteria are applied to patients treated with the relevant NOAC.

While RCTs remain the preferred approach for answering causal questions on efficacy and safety of different treatments, high-quality registries are critical for generating complementary evidence to support RCT results.[38] In particular, RCT emulation using real-world observational data may lead to a better understanding of how a treatment works in usual care settings versus the more constrained settings of RCTs. Moreover, evidence from observational registries comparing alternative treatments represents an important source of information about clinical endpoints for which RCTs are unavailable or unfeasible. Finally, whereas RCTs use randomisation to achieve comparability between trial arms, trial emulation requires statistical techniques to mimic randomised treatment allocation, as discussed in more detail below. As such methods require minimisation of unknown confounders for randomisation mimicking, high-quality registry data are essential to ensure the validity of trial emulation efforts.[24]

GARFIELD-AF provided this study with several unique strengths for trial emulation. First, unlike previous work which questioned the external validity of the landmark trials using retrospective data, our study used prospectively collected data with 2-year follow-up as in the original trials. Moreover, the registry allowed for assessment of all inclusion criteria, as well as the most important exclusion criteria of the landmark trials. Second, emulation of both trials using observational data was made feasible by application of appropriate methodology for non-randomised treatment allocation and confounder adjustment using propensity score overlap weighting.[23] Together, this resulted in one of the nearest possible comparisons of NOAC versus VKA in patients selected from a real-world registry. Finally, because NOACs emerged during the enrolment period of GARFIELD-AF, the registry is uniquely positioned to compare outcomes of their use with VKA use. Prior analyses suggested superior efficacy and safety of NOACs compared with VKA in newly diagnosed AF patients.[13] Our work now confirms that apixaban and rivaroxaban can be used safely in real-world patients conforming to ARISTOTLE and ROCKET AF trial inclusion criteria.

Our study had several limitations. First, although GARFIELD-AF has extensive baseline data, we were unable to assess all exclusion criteria defined by ARISTOTLE and ROCKET AF. Therefore, selection bias might have been introduced with the inclusion of AF patients who would have been excluded from the two target trials. Second, unlike randomised studies, observational studies are susceptible to unobserved confounding. This may partly explain the larger reduction in outcome risk and wider CIs than was seen in the landmark RCTs. To ensure comparability of treatment groups, we adjusted for an extensive list of confounding factors, though the possibility of unmeasured confounding cannot be ruled out. Third, treatment was defined as the first OAC received, meaning that our trial emulation was analogous to an intention-to-treat analysis. We were unable to account for non-recommended dosing, treatment switches or cessation, which prevented a per-protocol analysis. Fourth, while ROCKET AF restricted recruitment of low-risk patients, our analysis incorporated all ROCKET AF-eligible patients, resulting in a lower mean cardiovascular risk compared with the original trial ($CHADS_2$ score 2.6±0.8 vs. 3.5±1.0). We did not apply the same restrictions to the GARFIELD-AF dataset, as this would have further reduced the number of eligible rivaroxaban users for our analyses. Fifth, GARFIELD-AF recruited only newly diagnosed AF, whereas ARISTOTLE and ROCKET AF primarily recruited prevalent AF patients. While a number of aspects could be associated with this difference in type of recruited patients (eg, baseline characteristics, treatment choice or treatment (dis)continuation),[39] we could not determine how this may have affected the outcomes of our emulation study. Finally, our focus was on average patient characteristics and overall outcomes associated with RCT selection criteria. Generalisation of the results for real-world patient populations not meeting the criteria for these two trials was outside the scope of this work.[40]

CONCLUSIONS

The patient selection criteria of the ARISTOTLE and ROCKET AF trials limited the representativeness for real-world AF patients with newly diagnosed non-valvular AF.

Treatment comparisons based on the GARFIELD-AF observational data yielded results consistent with those of the two RCTs. Although RCTs remain the standard for comparing efficacy and safety of different treatments, our work indicates that, when using high-quality observational data and appropriate methodology for non-random treatment allocation, the emulation of target trials in real-world data can be successful.[23]

Author affiliations
[1]Thrombosis Research Institute, London, UK
[2]Department of General Practice, Amsterdam UMC Location AMC, Amsterdam, The Netherlands
[3]Amsterdam Public Health, Peronalized Medicine, Amsterdam, The Netherlands
[4]Cardiology Clinical Academic Group Molecular & Clinical Sciences Research Institute, St. George's University of London, London, UK
[5]Radboud Universiteit, Nijmegen, The Netherlands
[6]University of Besançon, Besançon, France
[7]Cyte Global, London, UK
[8]Department of Medicine and Geriatrics, Hospital Santa Marcelina, Sao Paulo, Brazil
[9]AZ Klina, General Hospital Klina, Brasschaat, Belgium
[10]General Medicine, Alfred Hospital, Monash University, Melbourne, Victoria, Australia

Acknowledgements We would like to thank the physicians, nurses and patients involved in the GARFIELD-AF registry. Editorial support was provided by Thomas Weissensteiner and Hamish Aitken-Buck (Thrombosis Research Institute, London, UK).

Collaborators A complete list of GARFIELD-AF investigators is given in online supplemental information.

Contributors All authors fulfilled the criteria under The International Committee of Medical Journal Editors Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals (ICMJE Recommendations 2019). JC, KP, FV, J-PB, FM, ACPB, FC, HG and AKK contributed to the design of the study and acquisition of the data. JCLH, SV and KP performed data analysis. JCLH drafted the manuscript and was the guarantor of this work. All authors were involved in editing and final approval of the manuscript and agreed to be accountable for the work.

Funding This work was supported by the Thrombosis Research Institute (London, UK). The GARFIELD-AF registry was sponsored by an unrestricted grant awarded by Bayer AG (Berlin, Germany) to the Thrombosis Research Institute.

Competing interests JC reports institutional grants and personal fees from Bayer, Boehringer Ingelheim, Pfizer/BMS and Daiichi Sankyo. KP has consultancies with Johnson & Johnson, Element Science, Artivion and Novartis. FV received grants from Bayer Healthcare and personal fees from Bayer Healthcare, BMS/Pfizer, Daiichi-Sankyo, and Boehringer-Ingelheim. J-PB received a personal fee from Thrombosis Research Institute, during the conduct of the study. FM is a former employee of Bayer AG. FC reports speaker fees from Boehringer Ingelheim Pharma, Bayer, Pfizer and Daiichi-Sankyo Europe and a modest research grant from Daiichi-Sankyo Europe. HG received honoraria from Bayer Australia, Eli Lilly Australia, Pfizer Australia and BMS Australia. AKK received personal fees and grants from Bayer AG, Sanofi S.A. and Anthos Therapeutics. JCLH, SV, REH and ACPB report no conflicts of interest.

Patient consent for publication Not applicable.

Ethics approval The registry protocol was approved by an independent ethics committee and hospital-based institutional review board. Individual study sites provided additional approvals. The registry was conducted in accordance with the Declaration of Helsinki, the International Conference on Harmonisation Good Pharmacoepidemiological and Clinical Practice Guidelines and local regulatory requirements. All study participants provided written informed consent. Anonymity and confidentiality of all participants are maintained.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Requests for patient-level data can be made to Saverio Virdone, head of statistics at the Thrombosis Research Institute (svirdone@tri-london.ac.uk). These requests should include a protocol summary and a summary of the statistical analysis plan. The request will be reviewed by the data sharing committee for approval and next steps will be discussed with the requestor.

ORCID iD

Jelle C L Himmelreich http://orcid.org/0000-0003-0430-1583

REFERENCES

1 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991;22:983–8.
2 Hindricks G, Potpara T, Dagres N. Corrigendum to: 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). *Eur Heart J* 2021;42:546–7.
3 Granger CB, Alexander JH, McMurray JJV, *et al*. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011;365:981–92.
4 Patel MR, Mahaffey KW, Garg J, *et al*. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med* 2011;365:883–91.
5 Connolly SJ, Ezekowitz MD, Yusuf S, *et al*. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009;361:1139–51.
6 Eisen A, Giugliano RP, Ruff CT, *et al*. Edoxaban vs warfarin in patients with nonvalvular atrial fibrillation in the US Food and Drug Administration approval population: An analysis from the Effective Anticoagulation with Factor Xa Next Generation in Atrial Fibrillation–Thrombolysis in Myocardial Infarction 48 (ENGAGE AF–TIMI 48) trial. *Am Heart J* 2016;172:144–51.
7 January CT, Wann LS, Calkins H, *et al*. 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2019;74:104–32.
8 Perry M, Kemmis Betty S, Downes N, *et al*. Atrial fibrillation: diagnosis and management—summary of NICE guidance. *BMJ* 2021;373:n1150.
9 Camm AJ, Accetta G, Ambrosio G, *et al*. Evolving antithrombotic treatment patterns for patients with newly diagnosed atrial fibrillation. *Heart* 2017;103:307–14.
10 Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?" *Lancet* 2005;365:82–93.
11 Lee S, Monz BU, Clemens A, *et al*. Representativeness of the dabigatran, apixaban and rivaroxaban clinical trial populations to real-world atrial fibrillation patients in the United Kingdom: a cross-sectional analysis using the General Practice Research Database. *BMJ Open* 2012;2:e001768.
12 Wedlund L, Kvedar J. Simulated trials: in silico approach adds depth and nuance to the RCT gold-standard. *NPJ Digit Med* 2021;4:121.
13 Camm AJ, Fox KAA, Virdone S, *et al*. Comparative effectiveness of oral anticoagulants in everyday practice. *Heart* 2021;107:962–70.
14 Villines TC, Peacock WF. Safety of Direct Oral Anticoagulants: Insights from Postmarketing Studies. *Am J Med* 2016;129:S41–6.
15 Kakkar AK, Mueller I, Bassand J-P, *et al*. International longitudinal registry of patients with atrial fibrillation at risk of stroke: Global Anticoagulant Registry in the FIELD (GARFIELD). *Am Heart J* 2012;163:13–9.
16 Fox KAA, Gersh BJ, Traore S, *et al*. Evolving quality standards for large-scale registries: the GARFIELD-AF experience. *Eur Heart J Qual Care Clin Outcomes* 2017;3:qcw058.
17 Lip GY, Nieuwlaat R, Pisters R, *et al*. 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;137:263–72.
18 Pisters R, Lane DA, Nieuwlaat R, *et al*. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. *Chest* 2010;138:1093–100.
19 Fox KAA, Virdone S, Pieper KS, *et al*. GARFIELD-AF risk score for mortality, stroke, and bleeding within 2 years in patients with atrial fibrillation. *Eur Heart J Qual Care Clin Outcomes* 2022;8:214–27.
20 Schulman S, Angerås U, Bergqvist D, *et al*. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in surgical patients. *J Thromb Haemost* 2010;8:202–4.
21 Gage BF, Waterman AD, Shannon W, *et al*. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. *JAMA* 2001;285:2864–70.
22 Azur MJ, Stuart EA, Frangakis C. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res* 2011;20:40–9.
23 Farjat AE, Virdone S, Thomas LE, *et al*. The importance of the design of observational studies in comparative effectiveness research: lessons from the GARFIELD-AF and ORBIT-AF registries. *Am Heart J* 2022;243:110–21.
24 Thomas LE, Li F, Pencina MJ. Overlap Weighting A Propensity Score Method That Mimics Attributes of a Randomized Clinical Trial. *J Am Med Assoc* 2020;323:2417–8.
25 Susukida R, Crum RM, Ebnesajjad C, *et al*. Generalizability of findings from randomized controlled trials: application to the National Institute of Drug Abuse Clinical Trials Network. *Addiction* 2017;112:1210–9.
26 Sam D, Gresham G, Abdel-Rahman O, *et al*. Generalizability of clinical trials of advanced melanoma in the real-world, population-based setting. *Med Oncol* 2018;35:110.
27 Leather DA, Jones R, Woodcock A, *et al*. Real-World Data and Randomised Controlled Trials: The Salford Lung Study. *Adv Ther* 2020;37:977–97.
28 Hughey AB, Gu X, Haymart B, *et al*. Warfarin for prevention of thromboembolism in atrial fibrillation: comparison of patient characteristics and outcomes of the "Real-World" Michigan Anticoagulation Quality Improvement Initiative (MAQI2) registry to the

RE-LY, ROCKET-AF, and ARISTOTLE trials. *J Thromb Thrombolysis* 2018;46:316–24.
29 Yoon CH, Park YK, Kim SJ, *et al*. Eligibility and preference of new oral anticoagulants in patients with atrial fibrillation: comparison between patients with versus without stroke. *Stroke* 2014;45:2983–8.
30 Hägg L, Johansson C, Jansson J-H, *et al*. External validity of the ARISTOTLE trial in real-life atrial fibrillation patients. *Cardiovasc Ther* 2014;32:214–8.
31 Desmaele S, Steurbaut S, Cornu P, *et al*. Clinical trials with direct oral anticoagulants for stroke prevention in atrial fibrillation: how representative are they for real life patients? *Eur J Clin Pharmacol* 2016;72:1125–34.
32 Fanning L, Ilomäki J, Bell JS, *et al*. The representativeness of direct oral anticoagulant clinical trials to hospitalized patients with atrial fibrillation. *Eur J Clin Pharmacol* 2017;73:1427–36.
33 Camm AJ, Amarenco P, Haas S, *et al*. Real-world vs. randomized trial outcomes in similar populations of rivaroxaban-treated patients with non-valvular atrial fibrillation in ROCKET AF and XANTUS. *Europace* 2019;21:421–7.
34 Laliberté F, Cloutier M, Nelson WW, *et al*. Real-world comparative effectiveness and safety of rivaroxaban and warfarin in nonvalvular atrial fibrillation patients. *Curr Med Res Opin* 2014;30:1317–25.
35 Amin A, Keshishian A, Trocio J, *et al*. Risk of stroke/systemic embolism, major bleeding and associated costs in non-valvular atrial fibrillation patients who initiated apixaban, dabigatran or rivaroxaban compared with warfarin in the United States Medicare population. *Curr Med Res Opin* 2017;33:1595–604.
36 Zhang C, Wang WW, Pan MM, *et al*. Simulation of anticoagulation in atrial fibrillation patients with rivaroxaban—from trial to target population. *Rev Cardiovasc Med* 2021;22:1019.
37 Himmelreich JCL, Virdone S, Camm J, *et al*. Impact of patient selection in clinical trials: application of ROCKET AF and ARISTOTLE criteria in GARFIELD-AF. *Open Heart* 2024;11:e002708.
38 Camm AJ, Fox KAA. Strengths and weaknesses of “real-world” studies involving non-vitamin K antagonist oral anticoagulants. *Open Heart* 2018;5:e000788.
39 Cools F, Johnson D, Camm AJ, *et al*. Risks associated with discontinuation of oral anticoagulation in newly diagnosed patients with atrial fibrillation: results from the GARFIELD-AF registry. *J Thromb Haemost* 2021;19:2322–34.
40 Laffin LJ, Besser SA, Alenghat FJ. A data-zone scoring system to assess the generalizability of clinical trial results to individual patients. *Eur J Prev Cardiol* 2019;26:569–75.