Simulation of anticoagulation in atrial fibrillation patients with rivaroxaban—from trial to target population

Chi Zhang[1,2,3,†], Wei-Wei Wang[4,†], Mang-Mang Pan[1], Zhi-Chun Gu[1,3,*]

[1] Department of Pharmacy, Ren Ji Hospital, Shanghai Jiao Tong University School of Medicine, 200127 Shanghai, China

[2] School of Medicine, Tongji University, 200092 Shanghai, China

[3] Shanghai Anticoagulation Pharmacist Alliance, Shanghai Pharmaceutical Association, 200040 Shanghai, China

[4] Department of S/4HANA Research & Development, SAP (China) Co., Ltd., 201203 Shanghai, China

*Correspondence: guzhichun213@163.com (Zhi-Chun Gu)
† These authors contributed equally.

The populations included in the randomized controlled clinical trials and observational studies were different. The effectiveness and safety of rivaroxaban for stroke prevention in patients with atrial fibrillation (AF) varied among studies. This study aimed to estimate the real-world outcomes of rivaroxaban in patients with AF accurately. A discrete event simulation (DES) was used to predict the counterfactual results of the ROCKET AF study. The hypothetical cohorts of patients were generated using Monte Carlo simulation according to the baseline covariate distributions that matched the marginal distribution of covariates reported in the ROCKET AF and three observational studies. The DES model structure was constructed based on *a priori* knowledge about disease progression and possible outcomes of patients with AF. The DES model accurately replicated the overall results of the ROCKET AF study. Both predicted stroke/systematic embolism (SE) and major bleeding rates were lower in the three observational studies than in the simulated ROCKET AF study. The risk difference of stroke/SE and major bleeding was not significant among the predicted outcomes of the three observational studies. Although some differences existed in the absolute rates of stroke/SE and major bleeding between observed and simulated studies, the results confirmed that rivaroxaban was noninferior to warfarin for the prevention of stroke/systematic embolism with no significance in the risk of major bleeding in large AF populations, which was similar to the results of ROCKET AF.

Keywords

Rivaroxaban; Atrial fibrillation (AF); Simulation; Randomized controlled trial (RCT); Observational study

1. Introduction

Atrial fibrillation (AF) is the most commonly diagnosed and treated arrhythmia in clinical practice, with an increasing health burden. In the United States, 2.7 to 6.1 million individuals are currently suffering from AF, and it is estimated to be prevalent in more than 8 million people by the year 2050 [1, 2]. Stroke is the most feared complication of AF, which is usually prevented by oral anticoagulation [3]. Non-vitamin K antagonist oral anticoagulants (NOACs) have become an alternative to vitamin K antagonists for preventing stroke in patients with AF [4]. Rivaroxaban, one of the most commonly used NOACs, was approved for stroke prevention in patients with AF based on the pivotal randomized controlled clinical trial (RCT), namely ROCKET AF study [5]. This trial demonstrated the efficacy and safety of rivaroxaban in reducing AF-related stroke risk.

RCTs are conducted on highly selective populations and are managed in tightly controlled settings. Therefore, RCTs are considered as the gold standard for assessing treatment efficacy, and the results have the highest reliability. Nevertheless, RCTs, such as the ROCKET AF study, usually exclude certain patient groups; for example, AF patients with a $CHADS_2$ score of 0 to 1 [5] account for about 40% of the entire AF population in the real-world setting [6, 7]. It is well known that AF patients with different stroke risks estimated by the $CHADS_2$ score might receive different benefits from the anticoagulation therapy [8]. Therefore, rigorous exclusions could limit the generalizability of evidence from RCTs, as the benefits and risks in all patient populations who actually treated in real clinical practice may differ. Consequently, uncertainty exists when physicians make anticoagulation decisions of rivaroxaban for patients who fall outside the inclusion criteria in the ROCKET AF study.

Real-world observational studies using routine electronic healthcare databases, such as insurance claims data or registry data, are available for large and diverse patient populations, and could be used to capture rare adverse events and long-term outcomes, as well as provide outcome estimates of treatment effectiveness in broad patient populations. However, the results of observational studies often differ from those of RCTs, which might mainly result from the differences in patient characteristics, drug adherence, and outcome measurement across studies that differ in design [9]. The XANTUS study, a real-world, prospective, observational cohort study described the use of rivaroxaban in a broad unselected AF pa-

tient population [6]. In both ROCKET AF and XANTUS studies, patients exhibited different baseline characteristics. In the XANTUS study, a lower CHADS$_2$ score and a lower proportion of patients with prior stroke, heart failure, hypertension, or diabetes mellitus compared to the ROCKET AF study were observed, which might have contributed to the outcome discrepancy between these two studies [10].

Moreover, the effectiveness and safety of rivaroxaban in patients with AF also varied among different observational studies. The incidence of stroke/systematic embolism (SE) was observed to be 0.8 event per 100 patient-year in the XANTUS study [6], whereas different rates (1.9 and 4.6 per 100 patient-year) were observed in two other observational studies [7, 11] (**Supplementary Table 1**). Such significant variations in observational studies might result from differences in study design, data source, definition of outcomes, length of observation, analysis methods, etc. [12]. Although real-world studies could support and extend RCT findings to larger patient populations, the results could be biased.

To obtain generalizable results on the use of rivaroxaban in patients with AF, it is desirable to make the results of the ROCKET AF study and real-world studies complement each other. To the best of our knowledge, generalizing the baseline characteristics of the RCT population to match those of real-world patients could facilitate the generation of evidence for effectiveness and safety of treatments in excluded populations, thus providing more relevant evidence for decision-makers [9]. It is worth noting that discrete event simulation (DES) is a method that can mimic the disease pathways and outcomes over time according to a function of treatment and patient-level covariates [13]. Using DES, we could generalize the results to real-world patients with a different distribution of characteristics from the RCT population. Previous studies [14, 15] have used DES to estimate the percentage of patients with atherosclerotic cardiovascular disease who would require lipid-lowering therapy (LLT) intensification in real practice and evaluate the impact of LLT intensification on cardiovascular events. In this study, we proposed a DES to predict the counterfactual outcomes of ROCKET AF that would have been conducted in larger observational study populations.

2. Methods

2.1 Data sources and study population

The ROCKET AF was selected as a case of RCT that evaluated the efficacy and safety of rivaroxaban versus warfarin in patients with AF. XANTUS, a prospective observational study, was chosen to investigate whether the results regarding the effectiveness and safety of rivaroxaban obtained in the ROCKET AF study could translate into real-world clinical practice. Two other retrospective observational studies (Laliberté, 2014 [7] and Amin, 2017 [11]) were used to assess the effectiveness and safety of rivaroxaban versus warfarin in routine care. Patient baseline characteristics of the four studies were collected, including age, sex, previous thromboembolic events (stroke, SE), transient ischaemic attack (TIA), myocardial infarction (MI), heart failure (HF), hypertension, diabetes mellitus (DM), and CHADS$_2$ risk of stroke. The incidence rate of events, such as stroke/SE, major bleeding, intracranial haemorrhage (ICH), gastrointestinal (GI) bleeding, and MI, were extracted as events per 100 patient-year.

2.2 Discrete event simulation (DES) model

A DES model was developed and reprogrammed in Python (version 3.7, The Python Software Foundation, USA). The hypothetical cohorts of patients were generated according to the baseline covariate distributions that matched the marginal distribution of covariates reported in the ROCKET AF, XANTUS, and two other observational studies (Laliberté, 2014 [7] and Amin, 2017 [11]), using Monte Carlo simulation by random sampling. Random sampling continued until 7000 patients were simulated for each treatment group, and this was similar to the sample size of the ROCKET AF study. Fig. 1 presents the DES model structure built based on *a priori* knowledge about disease progression and possible outcomes of patients with AF receiving rivaroxaban [3, 16]. The model was designed to predict treatment outcomes, based on patients' baseline characteristics. The CHADS$_2$ score of individual patients, representing their stroke risk, was calculated according to the simulated patients' baseline characteristics. Therefore, patients at different stroke risk would trace different probabilistic pathways in the model based on their treatment assignment (rivaroxaban or warfarin). The incidence of events in the simulation model was obtained based on the rates reported in the ROCKET AF study (**Supplementary Table 2**) [5, 17–21]. Although CHA$_2$DS$_2$-VASc score is now recommended in the clinical guidelines for stroke risk assessment of patients with AF, CHADS$_2$ score was the mainstream score for stroke prediction when the ROCKET AF study was conducted. Some patients with CHADS$_2$ score of 0 could obtain CHA$_2$DS$_2$-VASc score of 2 to 3, and those patients are eligible for anticoagulation therapy. As patients with a CHADS$_2$ score of 0–1 were excluded from the ROCKET AF study, the incidence of the events used in the model for this subset of patients was extrapolated based on the Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) trial, which investigated the efficacy and safety of dabigatran in AF patients. Cardiovascular events, such as stroke/SE, major bleeding (ICH, GI bleeding, etc.), non-major clinically relevant (NMCR) bleeding, MI, and unknown death, were recorded during the simulated two-year follow-up period. The patients' cardiovascular profiles were updated from the first year to the second year. Therefore, the CHADS$_2$ score of individual patients would be calculated again when entered the second year. Patients who suffered death, fatal stroke/SE, fatal major bleeding, or fatal MI, were removed from the simulation model after their outcomes were recorded. The event rates, hazard ratios (HRs), and risk differences (RDs) were calculated and reported.
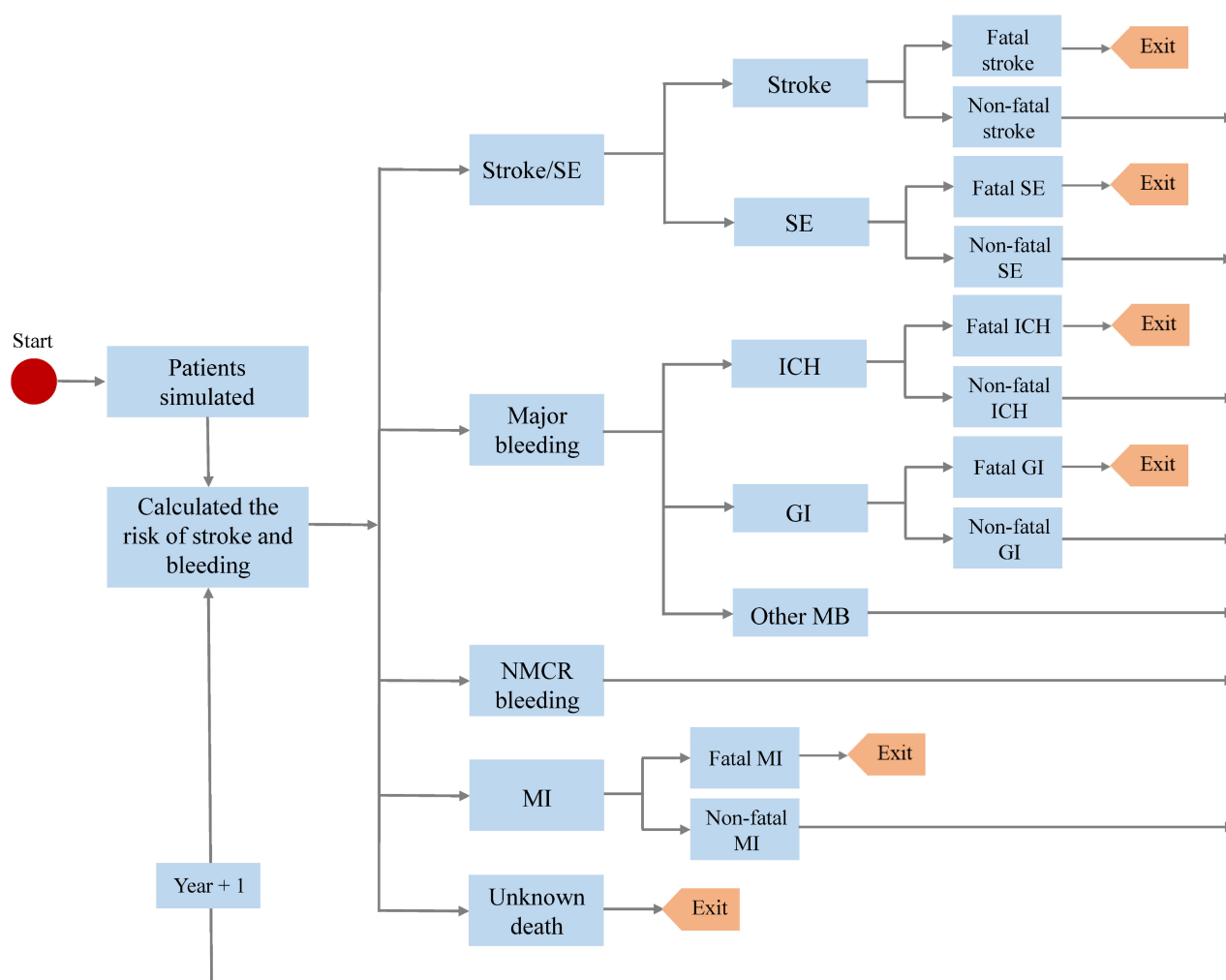
Fig. 1. The structure of discrete event simulation (DES) model. DES model structure was built based on *a priori* knowledge about disease progression and possible outcomes of AF patients. The model was designed to predict treatment outcomes conditional on patients' baseline characteristics. Patients at different stroke and major bleeding risk would trace different probabilistic pathways in the model based on their treatment assignment (rivaroxaban or warfarin). DES could keep track of patient-level covariates and account for the changes in patients' stroke and bleeding risk factors over time. Therefore, the stroke and major bleeding risk could be modified as the patient got older age or greater comorbidity burden.

2.3 Comparison of simulated and observed results in ROCKET AF

To validate the DES model, we compared the simulated results and observed results of the ROCKET AF study. The RDs and relative HRs (RHRs) of each outcome were calculated. RDs were estimated as the absolute risk difference of each outcome, and were calculated by subtracting the observed incidence from the simulated incidence [22]. HRs and 95% confidence intervals (CI) were calculated using the Cox proportional-hazards models. The RHRs were calculated by dividing the simulated HRs by the observed HRs for each outcome [22]. Both RDs and RHRs reflected the model error, which could have been caused by misspecification of the simulation structure or assumptions about input parameters. RDs around 0 ($\pm$0.10%) and RHRs near to1 (0.90–1.10) represented low simulation model error.

2.4 Comparison of simulated results among the ROCKET AF, XANTUS, and two observational studies

The simulated outcomes for hypothetical cohorts of patients with marginal covariate distributions similar to the XANTUS and two observational studies were predicted. Hypothetical cohorts of 7000 rivaroxaban patients and 7000 warfarin patients were simulated for the XANTUS and other two observational studies. As XANTUS was a single-arm study, and there was no baseline information about patients using warfarin; therefore, the baseline characteristics of patients using rivaroxaban were used in the simulation. The DES model developed and validated in patients with RCT was used to estimate outcomes of the three observational study populations. This process was conducted by replacing the baseline characteristics of the RCT cohorts with the simulated cohorts of the observational studies. The event rates, RDs, HRs, and RHRs were estimated.

Table 1. Baseline characteristics of rivaroxaban patients in ROCKET AF, XANTUS (2016) [6], Laliberté (2014) [7] and Amin (2017) [11].

Variables	ROCKET AF [5]	Laliberté (2014) [7]	Amin (2017) [11]	XANTUS (2016) [6]
No. of patient on rivaroxaban	7131	3654	52,467	6784
Age (year), mean $\pm$ SD	71.95 $\pm$ 9.64	73.3 $\pm$ 8.4	77.7 $\pm$ 7.2	71.5 $\pm$ 10
Female sex, no. (%)	2831 (39.7)	1865 (51.0)	27,135 (51.7)	2768 (40.8)
BMI, mean $\pm$ SD	28.54 $\pm$ 5.12	-	-	28.3 $\pm$ 5
Previous stroke, SE/TIA (%)	3916 (54.9)	357 (9.8)	6331 (12.1)	1291 (19)
HF, no. (%)	4467 (62.6)	716 (19.6)	15,239 (29)	1265 (18.6)
Hypertension, no. (%)	6436 (90.3)	2626 (71.9)	46,544 (88.7)	5065 (74.7)
DM, no. (%)	2878 (40.4)	919 (25.2)	18,989 (36.2)	1333 (19.6)
Previous MI, no. (%)	1182 (16.6)	-	6416 (12.2)	688 (10.1)
PAD, no. (%)	401 (5.6)	-	-	-
COPD, no. (%)	754 (10.6)	-	-	-
CHADS$_2$ risk of stroke				
CHADS$_2$ score, mean $\pm$ SD	3.48 $\pm$ 0.94	2 $\pm$ 1	2.7 $\pm$ 1.4	2 $\pm$ 1.3
Score, no. (%)				
0	0 (0)	0 (0)	1981 (3.8)	703 (10.4)
1	0 (0)	1464 (40.1)	9147 (17.4)	2061 (30.4)
2	925 (13)	1304 (35.7)	15,398 (29.3)	2035 (30)
3	3058 (42.9)	578 (15.8)		1111 (16.4)
4	2092 (29.3)	229 (6.3)	25,950 (49.5)	618 (9.1)
5	932 (13.1)	65 (1.8)		222 (3.3)
6	123 (1.7)	14 (0.4)		34 (0.5)

BMI, body-mess index; SE, systemic embolism; TIA, transient ischaemic attack; HF, heart failure; DM, diabetes mellitus; MI, myocardial infarction; PAD, peripheral vascular disease; COPD, chronic obstructive pulmonary disease; CHADS$_2$ score, the measure of the stroke risk ranges from 0 to 6.

2.5 Statistical analyses

Demographics and clinical characteristics at baseline for cohorts of involved studies and simulated cohorts are summarized descriptively as means $\pm$ SD or proportions, as appropriate. Key summary measures of the DES model were reported as event rates (events per 100 patient-year), RDs, HRs, and RHRs, with associated 95% CIs estimated using the 2.5th and 97.5th percentiles. Simulation of ROCKET AF was repeated 10, 20, 50, 100, 200, 300, 500, 1000, and 2000 times, respectively, to obtain the optimal iteration times for stable results. Finally, each simulation was repeated 1000 times with a computer running time of approximately 30 min. Analyses were performed using STATA software (version 13, StataCorp, College Station, Texas, USA) and Python (version 3.7, The Python Software Foundation, USA).

3. Results

3.1 Baseline characteristics of observed and simulated patients

Table 1 (Ref. [5–7, 11]) and **Supplementary Table 3** outline the baseline characteristics of the observed and simulated patients in the four studies. The average age in each of these four studies was above 70 years. The proportion of each comorbidity varied among the different studies. A larger proportion (54.9%) of patients on rivaroxaban with prior stroke/TIA was included in the ROCKET AF study than that in the three observational studies. Meanwhile, the proportion of patients with HF, hypertension, or DM was

also higher in the ROCKET AF study. The mean CHADS$_2$ score was 3.48 $\pm$ 0.94 in the ROCKET AF study, which was much higher than that in the three observational studies.

3.2 Simulation of ROCKET AF

The DES model accurately replicated the overall results of the ROCKET AF (Fig. 1 and Table 2). The simulation was repeated 1000 times to obtain robust and convergent results (**Supplementary Figs. 1,2**). The simulated incidence of stroke/SE and major bleeding was 1.718 vs. 1.980, and 3.463 vs. 3.379 per 100 patient-year for rivaroxaban and warfarin, respectively. The RDs between the simulated and observed results were relatively low among each outcome (Table 2). The simulated HRs comparing rivaroxaban and warfarin in the risks of stroke/SE and major bleeding were 0.868 (95% CI, 0.863–0.872) and 1.025 (95% CI, 1.021–1.029), respectively (Table 3, Ref. [5–7, 11]). The RHRs between the simulated and observed results were approximately 1. The results indicated that the estimated risks and HRs closely matched the observed risks and HRs in the ROCKET AF study, indicating a low error of the simulation model.

3.3 Counterfactual results of ROCKET AF using the populations in observational studies

Cohorts of equal size and similar covariate distributions were generated in the XANTUS, Laliberté (2014) [7] and Amin (2017) [11] studies, respectively (Table 1). The baseline characteristics of the ROCKET AF study were replaced

Table 2. The risk difference between observed and simulated results for ROCKET AF.

Events	Rivaroxaban			Warfarin		
	Observed rate	Simulated rate	Risk difference	Observed rate	Simulated rate	Risk difference
	no./100 patient-year	*no./100 patient-year*	*no./100 patient-year*	*no./100 patient-year*	*no./100 patient-year*	*no./100 patient-year*
Stroke/SE	1.7	1.718	0.018	2.2	1.980	−0.220
Stroke	1.65	1.677	0.027	1.96	1.804	−0.156
SE	0.04	0.042	0.002	0.19	0.175	−0.015
MB	3.6	3.463	−0.137	3.4	3.379	−0.021
ICH	0.5	0.484	−0.016	0.7	0.732	0.032
GIB	2	1.962	−0.038	1.24	1.352	0.112
MI	0.91	0.919	0.009	1.12	1.126	0.006

SE, systemic embolism; MB, major bleeding; ICH, intracranial hemorrhage; GIB, gastrointestinal bleeding; MI, myocardial infarction; RD, risk difference (simulated event minus observed event rate).

Table 3. Comparisons of each outcome between observed results and simulated results in ROCKET AF, XANTUS, and two observational studies.

Events	ROCKET AF [5]			Laliberté (2014) [7]			Amin (2017) [11]			XANTUS [6]
	Observed HR	Simulated HR	RHR	Observed HR	Simulated HR	RHR	Observed HR	Simulated HR	RHR	Simulated HR
	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)
Stroke/SE	0.79	0.868	1.099	0.77	0.780	1.013	0.72	0.824	1.144	0.737
	(0.65–0.95)	(0.863–0.872)	(0.904–1.335)	(0.55–1.09)	(0.775–0.785)	(0.715–1.435)	(0.63–0.83)	(0.819–0.829)	(0.991–1.322)	(0.732–0.742)
Stroke	0.85	0.930	1.094	-	0.835	-	0.70	0.883	1.261	0.789
	(0.70–1.03)	(0.924–0.935)	(0.897–1.335)		(0.829–0.840)		(0.59–0.83)	(0.878–0.889)	(1.057–1.505)	(0.784–0.794)
SE	0.23	0.253	1.100	-	0.233	-	0.52	0.257	0.494	0.221
	(0.09–0.61)	(0.246–0.260)	(0.411–2.944)		(0.225–0.242)		(0.28–0.94)	(0.249–0.266)	(0.261–0.936)	(0.213–0.229)
MB	1.04	1.025	0.986	1.08	0.940	0.870	1.17	1.034	0.884	0.899
	(0.90–1.20)	(1.021–1.029)	(0.850–1.142)	(0.71–1.64)	(0.936–0.944)	(0.570–1.328)	(1.10–1.26)	(1.030–1.039)	(0.822–0.950)	(0.895–0.903)
ICH	0.67	0.663	0.990	1.17	0.606	0.518	0.71	0.669	0.942	0.582
	(0.47–0.93)	(0.656–0.669)	(0.697–1.406)	(0.66–2.05)	(0.600–0.613)	(0.291–0.923)	(0.59–0.87)	(0.662–0.676)	(0.768–1.156)	(0.576–0.588)
GIB	1.66	1.450	0.873	1.27	1.333	1.050	1.35	1.468	1.087	1.278
	(1.34–2.05)	(1.442–1.459)	(0.702–1.087)	(0.99–1.63)	(1.324–1.341)	(0.813–1.355)	(1.23–1.48)	(1.460–1.477)	(0.986–1.200)	(1.27–1.286)
MI	0.81	0.817	1.009	-	0.813	-	-	0.811	-	0.807
	(0.63–1.06)	(0.811–0.823)	(0.772–1.318)		(0.807–0.819)			(0.805–0.816)		(0.801–0.813)

SE, systemic embolism; MB, major bleeding; ICH, intracranial hemorrhage; GIB, gastrointestinal bleeding; MI, myocardial infarction; HR, hazard ratio; 95% CI, 95% confidence interval; RHR, relative hazard ratio. RHR was calculated by dividing the simulated HR by observed HR. An RHR of 1 indicates no difference between simulated outcomes and observed outcomes.

with those of the simulated cohorts to repeat the simulation. The predicted outcomes are shown in Table 3, Table 4 (Ref. [5–7, 11]), and **Supplementary Tables 3,4,5**. The predicted rates of stroke/SE were 1.718, 1.118, 1.097, and 1.318 per 100 patient-year, while the predicted rates of major bleeding were 3.463, 2.817, 2.804, and 3.238 per 100 patient-year in rivaroxaban arms for simulated ROCKET AF, XANTUS, Laliberté (2014) [7], and Amin (2017) [11] studies, respectively. Both predicted rates of stroke/SE and major bleeding were lower in the three observational studies than those in the simulated ROCKET AF (Table 4), with RDs being 0.22–0.66 per 100 patient-year. However, the RDs of stroke/SE and major bleeding were similar among the predicted outcomes of the three observational studies, ranging from 0.02–0.43 per 100 patient-year (**Supplementary Tables 4,5**). Consistent effects were observed in other simulated outcomes such as stroke, ICH, GI bleeding, and MI. Considering the HRs be-

tween rivaroxaban and warfarin in each outcome, the simulated HRs of stroke/SE were 0.780 (95% CI, 0.775–0.785) and 0.824 (95% CI, 0.819–0.829) for Laliberté (2014) [7] and Amin (2017) [11] studies, respectively, which were close to the observed HRs of 0.77 (95% CI, 0.55–1.09) and 0.72 (95% CI, 0.63–0.83) (Table 3). The simulated HRs of major bleeding were 0.940 (95% CI, 0.936–0.944) for Laliberté (2014) [7] study and 1.034 (95% CI, 1.030–1.039) for Amin (2017) [11] study, which seemed to be relatively lower than the observed HRs of 1.08 (95% CI, 0.71–1.64) and 1.17 (95% CI, 1.10–1.26) for Laliberté and Amin, respectively. Besides, the simulated HRs of XANTUS study were similar with the observed HRs in ROCKET AF study, with relatively lower HR for GIB (1.278 [95% CI, 1.27–1.286]) detected in the simulation of XANTUS study (Table 3 and **Supplementary Table 6**). Even though some differences were observed between the event rates of each study, most of the simulated HRs were

Table 4. Simulated results and the risk difference for rivaroxaban arms in the simulation.

Events	ROCKET AF [5]	Laliberté (2014) [7]	Amin (2017) [11]	XANTUS (2016) [6]
Stroke/SE	1.718 (1.711–1.724)	1.097 (1.091–1.102)	1.318 (1.312–1.324)	1.118 (1.113–1.124)
RD (95% CI)	ref	0.62 (0.21–1.02)	0.40 (−0.02–0.82)	0.60 (0.20–1.01)
Stroke	1.677 (1.670–1.684)	1.07 (1.065–1.076)	1.287 (1.281–1.293)	1.092 (1.086–1.097)
RD (95% CI)	ref	0.61 (0.20–1.0)	0.39 (−0.03–0.80)	0.59 (0.19–0.99)
SE	0.042 (0.040–0.043)	0.026 (0.025–0.027)	0.032 (0.031–0.033)	0.026 (0.026–0.027)
RD (95% CI)	ref	0.02 (−0.08–0.11)	0.01 (−0.08–0.11)	0.02 (−0.08–0.11)
MB	3.463 (3.453–3.473)	2.804 (2.795–2.812)	3.238 (3.228–3.247)	2.817 (2.809–2.826)
RD (95% CI)	ref	0.66 (0.07–1.25)	0.22 (−0.39–0.82)	0.65 (0.05–1.23)
ICH	0.484 (0.480–0.487)	0.391 (0.388–0.395)	0.452 (0.449–0.456)	0.393 (0.390–0.397)
RD (95% CI)	ref	0.09 (−0.13–0.34)	0.03 (−0.21–0.27)	0.09 (−0.15–0.32)
GIB	1.962 (1.954–1.969)	1.585 (1.579–1.592)	1.84 (1.832–1.847)	1.603 (1.597–1.610)
RD (95% CI)	ref	0.38 (−0.08–0.82)	0.12 (−0.35–0.58)	0.36 (−0.09–0.81)
MI	0.919 (0.914–0.924)	0.913 (0.908–0.918)	0.914 (0.909–0.919)	0.912 (0.907–0.917)
RD (95% CI)	ref	0.01 (−0.33–0.33)	0.01 (−0.33–0.33)	0.01 (−0.33–0.33)

SE, systemic embolism; MB, major bleeding; ICH, intracranial hemorrhage; GIB, gastrointestinal bleeding; MI, myocardial infarction; RD, risk difference (simulated event rate of observational studies minus simulated event rate of ROCKET AF). The lower and upper limits of the 95% confidence interval for the RD between two studies were calculated using the website of http://vassarstats.net/ based on methods described by Robert Newcombe derived from a procedure outlined by E.B. Wilson in 1927.

similar to the corresponding observed HRs, with most RHRs around 1 (Table 3 and **Supplementary Table 6**).

4. Discussion

The ROCKET AF study and three observational studies (XANTUS, Laliberté (2014) [7] and Amin (2017) [11]) contributed to the clinical evidence for rivaroxaban in stroke prevention in patients with AF. However, the effectiveness and safety of rivaroxaban varied among the four studies. In this study, a DES model was proposed to predict the counterfactual outcomes of ROCKET AF study that could have been observed in larger observational study populations. The DES accurately replicated the overall results of the ROCKET AF study. Counterfactual results of the ROCKET AF study obtained by using the populations in observational studies showed relatively lower stroke/SE rates and major bleeding rates than those in the simulated ROCKET AF. Moreover, most of the simulated HRs between the rivaroxaban and warfarin arms were similar to the corresponding observed HRs,

indicating similarities in benefits and harms of rivaroxaban in patients with AF in the ROCKET AF study.

As an RCT, the ROCKET AF study investigating the efficacy and safety of rivaroxaban in AF patients is regarded as the gold standard. Nevertheless, the study was performed in selected AF patients with moderate-to-high risk of stroke (CHADS2 score $\geq$2 and mean score: 3.5), resulting in a lack of external validity and generalizability [23]. In comparison, real-world studies, such as the XANTUS, Laliberté (2014) [7], and Amin (2017) [11] studies, could reflect real-world treatment patterns among diverse populations and provide outcome estimates in broad patient populations. In fact, the results of observational studies often differ from those of RCTs and differ from each other. As baseline covariates, such as age and history of stroke, are also risk factors for the studied outcomes, differences in patient characteristics could be a common barrier, leading to outcome discrepancies across studies. As a result, the treatment effects might differ across different patient populations. In addition, differences in data

sources, outcome measures, and patient adherence as well as confounding bias of observational studies, contribute to the discrepant results. Furthermore, the follow-up periods were different in the XANTUS, Laliberté (2014) [7], and Amin (2017) [11] studies, ranging from 0.5 to 1 year, which was shorter than that in the ROCKET AF study (about 2 years). Considering that some rare events could not be detected and differences in low incidence events might not be found during the short follow-up period, the absolute event rates and the relative benefits of rivaroxaban might be inaccurate.

To estimate the real-world effectiveness and safety of rivaroxaban in patients with AF accurately, we used the DES method to model the pathways and 2-year outcomes of rivaroxaban anticoagulation in AF patients. It is known that DES is a strategy for modeling disease pathway and outcomes over time as a function of treatment and patient-level covariates [9]. Several studies have used DES to simulate LLT in patients with atherosclerotic cardiovascular disease [14, 15, 24]. Besides, a recent study used baseline characteristics from two observational studies to replicate the efficacy and safety of dabigatran compared to warfarin in the RE-LY study with a DES model [25]. The study found that differences in patient populations can explain a substantial portion of observed differences in outcomes across studies. In this study, a Monte Carlo simulation was used to generate hypothetical cohorts of patients. The DES built in this study could keep track of patient-level covariates and account for the changes in patients' stroke and bleeding risk factors over time [9]. Therefore, the stroke and major bleeding risks could be modified as the patient grew older or had a greater comorbidity burden. Event rates and treatment effects could then be estimated based on predefined relationships between the outcomes and risk factors of stroke and bleeding. The baseline characteristics of ROCKET AF patients were generalized to match the baseline of patients treated in routine care, which facilitated the generation of evidence for the effectiveness and safety of rivaroxaban in excluded AF populations.

Our results indicated that the observed outcomes of the ROCKET AF, XANTUS, Laliberté (2014) [7], and Amin (2017) [11] studies differed from each other. However, the differences were insignificant among the corresponding simulated studies. In the Laliberté (2014) [7] study, wide discrepancies were found in the stroke incidence in the rivaroxaban arm in both observed and simulated results, with an observed rate of 4.6 and a simulated rate of 1.097 per 100 patient-year. A similar trend was found in the observed and simulated incidence of major bleeding in the Amin (2017) [11] study. This inconsistency might be caused by the inherent limitations of real-world studies, such as short follow-up, unbalanced confounding bias, etc. Interestingly, stroke/SE incidence in the rivaroxaban group was similar in the simulated XANTUS, Laliberté (2014) [7], and Amin (2017) [11] studies (1.118, 1.097, and 1.318 per 100 patient-year, respectively), which was much lower than that in the simulated ROCKET AF study (1.718 per 100 patient-year). It is known that patients

enrolled in the ROCKET AF study were of moderate-to-high stroke risk. In comparison, the baseline characteristics of the three observational studies were similar and could represent the whole AF population. The stroke risk of patients in the three observational studies was much lower than that in the ROCKET AF study. Accordingly, the simulated stroke/SE incidence of the three observational studies might reflect the real-world stroke/SE rate in AF patients using rivaroxaban to some extent, same as the other simulated outcomes.

It is worth noting that most observed and simulated HRs between rivaroxaban and warfarin for each outcome were similar in our study, with most RHRs around 1. In terms of the HR for stroke/SE comparing rivaroxaban and warfarin, the simulated HRs in the Laliberté (2014) [7] and Amin (2017) [11] studies were 0.780 and 0.824, respectively, which were close to the observed HR of 0.79 in the ROCKET AF study. These results confirmed that rivaroxaban was non-inferior or even superior to warfarin for the prevention of stroke/SE in the real-world setting, which were in accordance with the results obtained in two previous meta-analyses reporting that HRs for stroke/SE comparing rivaroxaban and warfarin were 0.75 (95% CI, 0.64 to 0.85) and 0.83 (95% CI, 0.73 to 0.94) in real-world setting, respectively [26, 27]. With respect to the HR for major bleeding, there was no significant difference between groups, with observed HR being 1.04 in ROCKET AF study and simulated HR being 1.034 in Amin (2017) [11] study. The HRs for major bleeding obtained in this study were also similar to those reported in two previous meta-analyses considering real-world studies, with the HRs being 1.02 (95% CI, 0.95 to 1.10) and 0.99 (95% CI, 0.91 to 1.07), respectively [26, 27]. Therefore, similar results for the effectiveness and safety were observed when comparing rivaroxaban and warfarin in patients with AF.

It could be easier to understand the difference in effectiveness between rivaroxaban and warfarin when the molecular mechanisms of these two drugs are clarified. As a traditional oral anticoagulant, warfarin exerts its activity through inhibition of the synthesis of vitamin K dependent coagulation factors II, VII, IX, and X, as well as proteins C and S [28]. By comparison, rivaroxaban, as a NOAC, selectively inhibits free and clot-bound factor Xa, and further inhibits the generation of thrombin, thrombin-mediated activation of coagulation, and thrombin-mediated platelet aggregation [28]. It is notable that rivaroxaban can also exert an antiplatelet effect by acting through protease-activated receptor 1, which may lead to reduced frequency of atherothrombotic events and improved outcomes in patients [29].

This study has some limitations. First, the model error could not be neglected, as the DES model structure and pathway were built based on *a priori* knowledge regarding disease progression and possible outcomes of AF patients receiving rivaroxaban, which lacked a multivariable outcome prediction component. Second, the CHA$_2$DS$_2$-VASc score, rather than the CHADS$_2$ score, is now recommended in the clinical guidelines for stroke risk assessment of patients with AF, as

it has the advantage to identify a subset of low-risk AF patients with CHADS$_2$ score of 0–1 [8, 30]. However, in this study, the relationship between baseline characteristics and clinical outcomes was calculated according to the patient's CHADS$_2$ score, as it was the mainstream score for stroke prediction when the ROCKET AF study was conducted. Third, as the ROCKET AF study excluded patients with a CHADS$_2$ score of 0 to 1, the incidence of the events used in the simulation model for this subset of patients was based on the RE-LY trial, which investigated the efficacy and safety of dabigatran in patients with AF. This could also introduce some errors in the model. Moreover, individual-level information was not available in our study. Therefore, the bootstrapping method, which could preserve the covariance structure among the baseline characteristics of observational studies and could increase the accuracy of the simulation, could not be used. In addition, the covariates and outcomes of observational studies might be imprecise, as the data were not originally recorded for research purposes and some vital information might be missing. All these factors might have negatively impacted the accuracy of the simulation model and the predicted outcomes.

5. Conclusions

To estimate the real-world effectiveness and safety of rivaroxaban in patients with AF, the DES method was used to model the pathways and 2-year outcomes of rivaroxaban anticoagulation in these patients. The simulated event incidence of observational studies, such as stroke/SE incidence and major bleeding incidence, which was lower than that in the simulated ROCKET AF study, might reflect the real-world event rate in patients with AF. Moreover, the results confirmed that rivaroxaban was noninferior to warfarin for prevention of stroke/SE with no significance in the risk of major bleeding in large AF populations, which was similar to the results of ROCKET AF study.

Author contributions

ZCG and CZ designed the study. MMP and WWW collected and analyzed the data. WWW was responsible for methodology and software. CZ wrote the original manuscript. All authors have read and agreed to the published version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Conflict of interest

The authors declare no conflict of interest.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at https://rcm.imrpress.com/EN/10.31083/j.rcm2203111.

Data availability

The data that support the findings of this study are available on reasonable request from the corresponding author.

References

[1] Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of Current and Future Incidence and Prevalence of Atrial Fibrillation in the U.S. Adult Population. The American Journal of Cardiology. 2013; 112: 1142–1147.

[2] Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, et al. Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. Circulation. 2004; 110: 1042–1046.

[3] January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, 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 in Collaboration With the Society of Thoracic Surgeons. Circulation. 2019; 140: e125–e151.

[4] Steffel J, Verhamme P, Potpara TS, Albaladejo P, Antz M, Desteghe L, et al. The 2018 European Heart Rhythm Association Practical Guide on the use of non-vitamin K antagonist oral anticoagulants in patients with atrial fibrillation. European Heart Journal. 2018; 39: 1330–1393.

[5] Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. The New England Journal of Medicine. 2011; 365: 883–891.

[6] Camm AJ, Amarenco P, Haas S, Hess S, Kirchhof P, Kuhls S, et al. XANTUS: a real-world, prospective, observational study of patients treated with rivaroxaban for stroke prevention in atrial fibrillation. European Heart Journal. 2016; 37: 1145–1153.

[7] Laliberté F, Cloutier M, Nelson WW, Coleman CI, Pilon D, Olson WH, et al. Real-world comparative effectiveness and safety of rivaroxaban and warfarin in nonvalvular atrial fibrillation patients. Current Medical Research and Opinion. 2014; 30: 1317–1325.

[8] Lip GYH. Stroke and bleeding risk assessment in atrial fibrillation: when, how, and why? European Heart Journal. 2013; 34: 1041–1049.

[9] Wang SV, Schneeweiss S, Gagne JJ, Evers T, Gerlinger C, Desai R, et al. Using Real-World Data to Extrapolate Evidence from Randomized Controlled Trials. Clinical Pharmacology & Therapeutics. 2019; 105: 1156–1163.

[10] Camm AJ, Amarenco P, Haas S, Hess S, Kirchhof P, Lambelet M, 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. EP Europace. 2019; 21: 421–427.

[11] Amin A, Keshishian A, Trocio J, Dina O, Le H, Rosenblatt L, *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. Current Medical Research and Opinion. 2017; 33: 1595–1604.

[12] Blonde L, Khunti K, Harris SB, Meizinger C, Skolnik NS. Interpretation and Impact of Real-World Clinical Data for the Practicing Clinician. Advances in Therapy. 2018; 35: 1763–1774.

[13] Caro JJ, Möller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? Value in Health. 2010; 13: 1056–1060.

[14] Cannon CP, Khan I, Klimchak AC, Reynolds MR, Sanchez RJ, Sasiela WJ. Simulation of Lipid-Lowering Therapy Intensification in a Population with Atherosclerotic Cardiovascular Disease. JAMA Cardiology. 2017; 2: 959–966.

[15] Cannon CP, Khan I, Klimchak AC, Sanchez RJ, Sasiela WJ, Massaro JM, *et al*. Simulation of impact on cardiovascular events due to lipid-lowering therapy intensification in a population with atherosclerotic cardiovascular disease. American Heart Journal. 2019; 216: 30–41.

[16] Kim H, Kim H, Cho SK, Kim JB, Joung B, Kim C. Cost-Effectiveness of Rivaroxaban Compared to Warfarin for Stroke Prevention in Atrial Fibrillation. Korean Circulation Journal. 2019; 49: 252–263.

[17] Oldgren J, Alings M, Darius H, Diener H, Eikelboom J, Ezekowitz MD, *et al*. Risks for stroke, bleeding, and death in patients with atrial fibrillation receiving dabigatran or warfarin in relation to the CHADS_2 score: a subgroup analysis of the RE-LY trial. Annals of Internal Medicine. 2011; 155: 660–667.

[18] Goodman SG, Wojdyla DM, Piccini JP, White HD, Paolini JF, Nessel CC, *et al*. Factors associated with major bleeding events: insights from the ROCKET AF trial (rivaroxaban once-daily oral direct factor Xa inhibition compared with vitamin K antagonism for prevention of stroke and embolism trial in atrial fibrillation). Journal of the American College of Cardiology. 2014; 63: 891–900.

[19] Mahaffey KW, Stevens SR, White HD, Nessel CC, Goodman SG, Piccini JP, *et al*. Ischaemic cardiac outcomes in patients with atrial fibrillation treated with vitamin K antagonism or factor Xa inhibition: results from the ROCKET AF trial. European Heart Journal. 2014; 35: 233–241.

[20] Hankey GJ, Stevens SR, Piccini JP, Lokhnygina Y, Mahaffey KW, Halperin JL, *et al*. Intracranial Hemorrhage among Patients with Atrial Fibrillation Anticoagulated with Warfarin or Rivaroxaban. Stroke. 2014; 45: 1304–1312.

[21] Sherwood MW, Nessel CC, Hellkamp AS, Mahaffey KW, Piccini JP, Suh E, *et al*. Gastrointestinal Bleeding in Patients with Atrial Fibrillation Treated with Rivaroxaban or Warfarin: ROCKET AF Trial. Journal of the American College of Cardiology. 2015; 66: 2271–2281.

[22] Andrade C. Understanding relative risk, odds ratio, and related terms: as simple as it can get. The Journal of Clinical Psychiatry. 2015; 76: e857–e861.

[23] Camm AJ, Coleman CI, Larsen TB, Nielsen PB, Tamayo CS. Understanding the Value of Real-World Evidence: Focus on Stroke Prevention in Atrial Fibrillation with Rivaroxaban. Thrombosis and Haemostasis. 2018; 118: S45–S60.

[24] Allahyari A, Jernberg T, Hagström E, Leosdottir M, Lundman P, Ueda P. Application of the 2019 ESC/EAS dyslipidaemia guidelines to nationwide data of patients with a recent myocardial infarction: a simulation study. European Heart Journal. 2020; 41: 3900–3909.

[25] Najafzadeh M, Schneeweiss S, Choudhry NK, Wang SV, Gagne JJ. Simulation for Predicting Effectiveness and Safety of New Cardiovascular Drugs in Routine Care Populations. Clinical Pharmacology & Therapeutics. 2018; 104: 1008–1015.

[26] Escobar C, Martí-Almor J, Pérez Cabeza A, Martínez-Zapata MJ. Direct Oral Anticoagulants Versus Vitamin K Antagonists in Real-life Patients with Atrial Fibrillation. a Systematic Review and Meta-analysis. Revista Espanola De Cardiologia. 2019; 72: 305–316.

[27] Bai Y, Deng H, Shantsila A, Lip GYH. Rivaroxaban Versus Dabigatran or Warfarin in Real-World Studies of Stroke Prevention in Atrial Fibrillation: Systematic Review and Meta-Analysis. Stroke. 2017; 48: 970–976.

[28] DeWald TA, Washam JB, Becker RC. Anticoagulants: Pharmacokinetics, Mechanisms of Action, and Indications. Neurosurgery Clinics of North America. 2018; 29: 503–515.

[29] Petzold T, Thienel M, Dannenberg L, Mourikis P, Helten C, Ayhan A, *et al*. Rivaroxaban Reduces Arterial Thrombosis by Inhibition of FXa-Driven Platelet Activation via Protease Activated Receptor-1. Circulation Research. 2020; 126: 486–500.

[30] Chen JY, Zhang AD, Lu HY, Guo J, Wang FF, Li ZC. CHADS_2 versus CHA_2DS_2-VASc score in assessing the stroke and thromboembolism risk stratification in patients with atrial fibrillation: a systematic review and meta-analysis. Journal of Geriatric Cardiology. 2013; 10: 258–266.