ESC European Society of Cardiology

European Heart Journal - Digital Health (2024) **5**, 235–246
https://doi.org/10.1093/ehjdh/ztae010

ORIGINAL ARTICLE
Artificial intelligence (machine learning, deep learning)

Predicting multifaceted risks using machine learning in atrial fibrillation: insights from GLORIA-AF study

Juan Lu [1,2,3,4,†], Arnaud Bisson[1,5,†], Mohammed Bennamoun[3], Yalin Zheng[1,6], Frank M. Sanfilippo [7], Joseph Hung [2], Tom Briffa[7], Brendan McQuillan[2,8], Jonathon Stewart[2], Gemma Figtree[9,10], Menno V. Huisman[11], Girish Dwivedi [2,3,12,*‡], and Gregory Y.H. Lip [1,13,*‡]

[1]Liverpool Centre for Cardiovascular Science at University of Liverpool, Liverpool John Moores University and Liverpool Heart & Chest Hospital, Thomas Drive, Liverpool L14 3PE, UK; [2]Medical School, The University of Western Australia, 35 Stirling Hwy, Crawley WA 6009, Australia; [3]Harry Perkins Institute of Medical Research, 5 Robin Warren Dr, Murdoch WA 6150, Australia; [4]Department of Computer Science and Software Engineering, The University of Western Australia, 35 Stirling Hwy, Crawley WA 6009, Australia; [5]Department of Cardiology, University Hospital and University of Tours, Tours, France; [6]Department of Eye and Vision Sciences, University of Liverpool, Liverpool, UK; [7]School of Population and Global Health, The University of Western Australia, Perth, Australia; [8]Sir Charles Gairdner Hospital, Perth, Australia; [9]Kolling Institute and Charles Perkins Centre, University of Sydney, Sydney, Australia; [10]Department of Cardiology, Royal North Shore Hospital, Sydney, Australia; [11]Department of Thrombosis and Hemostasis Leiden University Medical Center, Leiden, The Netherlands; [12]Department of Cardiology, Fiona Stanley Hospital, Perth, Australia; and [13]Danish Center for Health Services Research, Department of Clinical Medicine, Aalborg University, Selma Lagerløfs Vej 249, 9260 Gistrup, Denmark

Received 29 August 2023; revised 22 December 2023; accepted 29 December 2023; online publish-ahead-of-print 20 February 2024

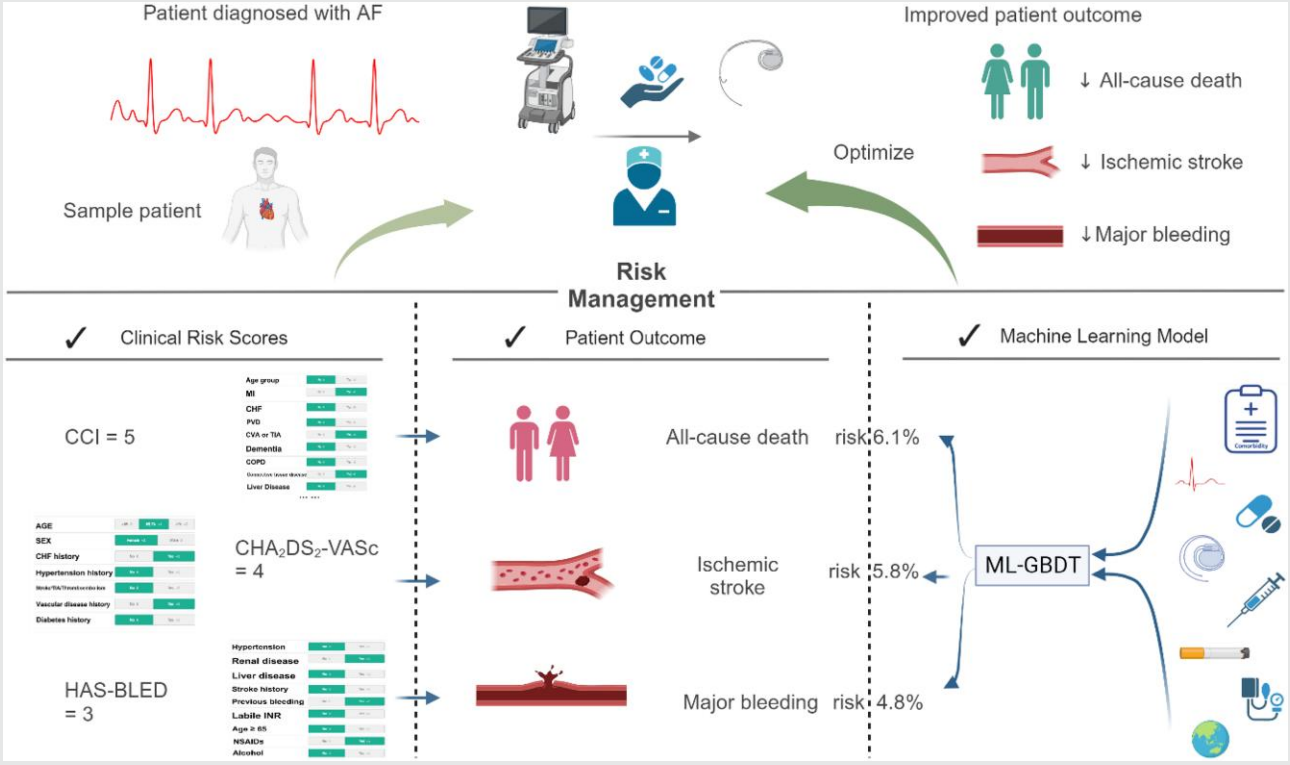
Aims Patients with atrial fibrillation (AF) have a higher risk of ischaemic stroke and death. While anticoagulants are effective at reducing these risks, they increase the risk of bleeding. Current clinical risk scores only perform modestly in predicting adverse outcomes, especially for the outcome of death. We aimed to test the multi-label gradient boosting decision tree (ML-GBDT) model in predicting risks for adverse outcomes in a prospective global AF registry.

Methods and results We studied patients from phase II/III of the Global Registry on Long-Term Oral Anti-Thrombotic Treatment in Patients with Atrial Fibrillation registry between 2011 and 2020. The outcomes were all-cause death, ischaemic stroke, and major bleeding within 1 year following the AF. We trained the ML-GBDT model and compared its discrimination with the clinical scores in predicting patient outcomes. A total of 25 656 patients were included [mean age 70.3 years (SD 10.3); 44.8% female]. Within 1 year after AF, ischaemic stroke occurred in 215 (0.8%), major bleeding in 405 (1.6%), and death in 897 (3.5%) patients. Our model achieved an optimized area under the curve in predicting death (0.785, 95% CI: 0.757–0.813) compared with the Charlson Comorbidity Index (0.747, $P = 0.007$), ischaemic stroke (0.691, 0.626–0.756) compared with CHA_2DS_2-VASc (0.613, $P = 0.028$), and major bleeding (0.698, 0.651–0.745) as opposed to HAS-BLED (0.607, $P = 0.002$), with improvement in net reclassification index (10.0, 12.5, and 23.6%, respectively).

Conclusion The ML-GBDT model outperformed clinical risk scores in predicting the risks in patients with AF. This approach could be used as a single multifaceted holistic tool to optimize patient risk assessment and mitigate adverse outcomes when managing AF.

* Corresponding author. Tel: +61 8 6488 8181, Email: girish.dwivedi@perkins.uwa.edu.au (G.D.); Tel: +44 151 795 1363, Email: gregory.lip@liverpool.ac.uk (G.Y.H.L.)
†Joint first authors.
‡Joint senior authors.
Drs Lip and Huisman are co-Chairs of the GLORIA-AF registry.

Graphical Abstract



Currently, clinical risk scores, derived from a specific set of weighted risk factors, are utilized to predict risks of all-cause mortality, ischaemic stroke, and major bleeding for patients diagnosed with AF. Consider a patient as an example: a patient might score 5 on the CCI, 4 on the CHA_2DS_2-VASc, and 3 on the HAS-BLED. Based on these scores, clinicians use predefined cut-offs to determine the suitable treatments. We aim to build a machine learning model, which integrates additional data (both input and output information) to improve risk prediction. Clinicians could leverage the developed model to enhance the overall outcomes for AF patients, benefiting from its superior predictive capabilities and the clarity provided by its feature importance ranking.

AF, atrial fibrillation; CCI, Charlson Comorbidity Index; CHA_2DS_2-VASc includes congestive heart failure, hypertension, age, diabetes mellitus, stroke or transient ischaemic attack, vascular disease, and sex (female); HAS-BLED includes hypertension, abnormal renal or liver function, stroke, bleeding, international normalized ratio, elderly (age over 65 years), and drugs or alcohol usage; ML-GBDT, multi-label gradient boosting decision tree.

Keywords Atrial fibrillation • Death • Ischaemic stroke • Major bleeding • Machine learning • Risk

Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia, increasingly affecting over 45 million people globally.[1] Patients diagnosed with AF have significantly elevated risks of ischaemic stroke and death.[2] Oral anticoagulation therapy is used to mitigate the risks of ischaemic stroke and all-cause death, but it increases the risk of major bleeding.[3,4] It is also well recognized that patients may still suffer from residual risks of death and ischaemic stroke despite receiving anticoagulants.[5,6]

The CHA_2DS_2-VASc[7] and HAS-BLED[8] scores are recommended by clinical guidelines for evaluating stroke and bleeding risks, respectively.[9,10] However, these scores have only a modest discriminatory capacity for ischaemic stroke and major bleeding in patients with AF.[11,12] Apart from the risk factors employed in these simple clinical risk scores, ethnicity or geographic regions have been identified as potential additional factors that can influence the prognosis for patients after being diagnosed with AF.[13] For example, studies have shown that the modified risk stratification tool mCHA_2DS_2-VASc tends to provide more accurate risk predictions for Asian patients with AF.[14] Risk factors for stroke or bleeding are also risks for death, although the Charlson Comorbidity Index (CCI)[15] is an established tool applied to predict one-year death based on specific comorbidities.[16]

Machine learning (ML) algorithms have shown better risk-predictive performance, after accounting for the presence of multimorbid conditions and dynamic changes in risk.[17] Additionally, multi-label ML can predict different patient outcomes simultaneously (e.g. stroke, major bleeding, and death) taking into account the interdependence between risks, therefore offering a more effective and multifaceted approach to predicting event risks faced by patients with AF. This would allow us to predict death, as well as ischaemic stroke and major bleeding risks for patients with AF by incorporating their diverse comorbidities, existing treatment strategies, and demographics including ethnicity and geographical region. This offers valuable insights for early intervention on the most significant risks, leading to improved clinical outcomes.

In our previous study in an Australian AF registry, we constructed and assessed different multi-label ML algorithms, among which the multi-label gradient boosting decision tree (ML-GBDT) model demonstrated superior performance for predicting stroke and bleeding risk in patients with non-valvular AF compared with clinical risk scores (CHA_2DS_2-VASc and HAS-BLED).[18] One major limitation of this

previous study was the retrospective nature of the dataset which was sourced exclusively from one geographical region with the predominant ethnic group being Caucasians. Moreover, non-vitamin K antagonist oral anticoagulants (NOACs) had not yet reached widespread use in our previous study. An external validation of our approach in global prospectively collected AF cohorts would be highly desirable.

Herein, we aimed to investigate our prior ML-GBDT algorithms to predict death, ischaemic stroke, and major bleeding within the first year after the baseline visit for patients with newly diagnosed non-valvular AF. We used a large, contemporary global cohort of patients from the prospective GLORIA-AF (Global Registry on Long-Term Oral Anti-Thrombotic Treatment in Patients with Atrial Fibrillation) registry.[19] Our goal was to assess the applicability and utility of multi-label ML across diverse AF patient populations and determine its potential for widespread clinical adoption.

Methods

We used data from Phase II/III of the prospective GLORIA-AF registry, which is a prospective international registry that gathered information about patients from 935 medical centres in 38 countries across Asia, Europe, North America, Latin America, and Africa/Middle East.[19] The registry comprised adults with newly diagnosed non-valvular AF (within 3 months prior to baseline visit) and an increased risk of stroke (CHA_2DS_2-VASc score ≥ 1).

We included patients (all five regions) from the GLORIA-AF registry enrolled between 2011 and 2020 who had at least 1 year of follow-up or died during the first year. We formed subsets based on age (age 65 years or more), anticoagulant treatment, residential region (Europe, North America, and Asia, excluded Latin America and Africa/Middle East due to the limited data samples collected from these regions), and whether patients had primary prevention or secondary prevention of ischaemic stroke. This enabled an in-depth exploration into the performance of different risk prediction models within the sub-cohorts sharing certain characteristics, and thereby providing nuanced insights into their predictive capacities, identifying factors that may impact the risk prediction in patients with AF.

AF classification was as per the European Society of Cardiology recommendations.[20] The severity of AF-related symptoms was determined based on the European Heart Rhythm Association classification.[21] The CHA_2DS_2-VASc, HAS-BLED scores, and CCI were calculated from data within the registry. Due to the limitations in data availability, specifically for comorbidities such as lymphoma and acquired immunodeficiency syndrome, we employed a tailed adaptation of the CCI.

Predictors

We included 110 variables (features) in the initial model, including patient demographics, vital signs, comorbidity history, medication treatment, medical procedures, and laboratory results. The complete list of variables is available in Supplementary material online, *Table S1*. We did not include variables with more than 25% missing values ($n = 16$) and imputed the remaining missing values using the multivariate feature imputation method.[22] Detailed missing rate information is presented in Supplementary material online, *Table S2*. Categorical variables were presented as counts and percentages and tested for differences with the χ^2 test. Continuous variables are presented as mean ± standard deviation (SD) and tested for differences with the *t*-tests. A two-sided $P < 0.05$ was considered statistically significant.

Outcome

Follow-up was 1 year from the baseline AF diagnosis. The primary outcome of interest was all-cause death, ischaemic stroke, and major bleeding within the 1-year follow-up. Additionally, we also investigated the short-term risks associated with these outcomes, specifically within the first 90 days. Stroke was defined as an acute onset of a localized neurological impairment which originated from a vascular source and was continuous for a minimum of 24 h or led to death. Data from computed tomography scans, magnetic resonance imaging, or autopsy were applied by the GLORIA-AF investigators to determine the presence of ischaemic stroke.[6] Major bleeding was classified by the GLORIA-AF investigators according to the definition from the International Society on Thrombosis and Hemostasis.[23]

Machine learning model

A gradient boosting decision tree[19]-based classifier chain,[20] namely, ML-GBDT, was the best-performing algorithm (compared with support vector machine and multi-layer neural networks) in previous studies for predicting risks in AF patients.[18] The order of the classifier chain was established as major bleeding, ischaemic stroke, and all-cause death based on our previous experimental results[18] and time-to-event for different outcomes. The predicted probability of major bleeding was used to predict ischaemic stroke, and both predictions were used to predict all-cause death. This effectively empowered the model to handle patients with multiple events. For comparison purposes, we also constructed single-label GBDT models to predict each outcome independently. Modelling and statistical analyses were performed using Python 3.9 and Stata (16.1, StataCorp LLC, USA).

Development vs. validation

Compared with our previous study, predictors were modified based on the available variables.[18] We added medical treatments for AF as predictors but were unable to incorporate predictors such as haemoglobin and indigenous ethnicity, which were not recorded in the GLORIA-AF registry. We utilized different data sources to define outcomes. In the current study, we used clinical information. We also implemented different hyperparameter tuning strategies while doing the validation. The feature importance measurement strategy was updated to permutation features importance, which tends not to be biased on high cardinality features.

Sampling and processing

We split the data into three sets for different purposes. We randomly allocated 70% of the data for training and internal validation; 10% of the training data were used for internal validation and hyperparameter tuning. The remaining 30% of the data were assigned for testing. Text data were encoded as binary or categorical variables. The continuous variables were standardized with a mean of 0 and a SD of 1, ensuring that all continuous features are on a similar scale.[24]

Feature importance

We included all the relevant available variables in the initial model. We then used the permutation feature importance[25] and presented 25–30 top-ranked features using bar plots. The permutation feature importance was defined as the reduction in a model's performance metrics when the values of an individual feature were randomly shuffled. The feature importance ranking plot was presented as a horizontal bar chart, with each bar representing a variable in our data. The length of the bar indicated the level of the impact on the ML-GBDT predictive performance.

Performance measurement

Measures of prediction performance included the area under the receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, and net reclassification index (NRI). A probability threshold of 0.5 was selected to determine the sensitivity and specificity of clinical risk scores. We then reserved a similar specificity for the ML-GBDT model while choosing the thresholds and calculated the sensitivity. A decision curve analysis was employed to show the net benefits of each model across a range of threshold probabilities, and Kaplan–Meier survival curves were presented to visualize the probability of an event occurring over the 3-year follow-up time based on the predicted probability.

Application

We developed an online risk calculator for clinical use, selected only 25 key variables through permutation feature importance predicting all-cause death, ischaemic stroke, and major bleeding risks within the 1-year follow-up. Despite this simplification, we managed to maintain a similar AUC as the model using the full variable set. The primary goal of this tool is to equip healthcare practitioners with the means to provide personalized, refined risk assessments for patients with AF.

Table 1 Baseline characteristics of the study cohort identified from Phase II/III of the Global Registry on Long-Term Oral Anti-Thrombotic Treatment in Patients with Atrial Fibrillation registry stratified by outcome

	All-cause death			Ischaemic stroke			Major bleeding		
	No event ($n = 24\,759$)	Event ($n = 897$)	*P*-value	No event ($n = 25\,441$)	Event ($n = 215$)	*P*-value	No event ($n = 25\,251$)	Event ($n = 405$)	*P*-value
Demographics									
Age, mean (SD)	70.1 (10.3)	76.2 (9.2)	<0.0001	70.2 (10.3)	74.5 (9.2)	<0.0001	70.2 (10.3)	74.3 (8.9)	<0.0001
Female, No. (%)	11 114 (44.9)	369 (41.1)	0.03	11 387 (44.8)	96 (44.7)	0.97	11 299 (44.7)	184 (45.4)	0.78
BMI (kg/m^2)	28.7 (6.2)	27.1 (6.6)	<0.0001	28.6 (6.3)	27.2 (5.6)	0.001	28.6 (6.3)	28.6 (6.2)	0.8
Region, No. (%)									
Europe	12 325 (49.8)	467 (52.1)	<0.0001	12 684 (49.9)	108 (50.2)	0.066	12 593 (49.9)	199 (49.1)	<0.0001
North America	5593 (22.6)	201 (22.4)		5754 (22.6)	40 (18.6)		5651 (22.4)	143 (35.3)	
Asia	4661 (18.8)	108 (12.0)		4715 (18.5)	54 (25.1)		4726 (18.7)	43 (10.6)	
Latin America	1874 (7.6)	115 (12.8)		1977 (7.8)	12 (5.6)		1969 (7.8)	20 (4.9)	
Africa/Middle East	306 (1.2)	6 (0.7)		311 (1.2)	1 (0.5)		312 (1.24)	0 (0.0)	
Race, No. (%)									
White	18 840 (76.1)	716 (79.8)	<0.0001	19 405 (76.3)	151 (70.2)	0.023	19 221 (76.1)	335 (82.7)	<0.0001
Asian	4333 (17.5)	112 (12.5)		4393 (17.3)	52 (24.2)		4399 (17.4)	46 (11.4)	
Black or Afro-Caribbean	424 (1.7)	22 (2.5)		440 (1.7)	6 (2.8)		431 (1.7)	15 (3.70)	
Arab or Middle East	324 (1.3)	5 (0.56)		329 (1.3)	0 (0.0)		327 (1.3)	2 (0.49)	
Other	838 (3.4)	42 (4.7)		874 (3.4)	6 (2.8)		873 (3.5)	7 (1.73	
Vital signs									
Heart rate (bpm)	80.0 (21.2)	83.7 (20.6)	<0.0001	80.1 (21.2)	79.7 (18.7)	0.77	80.1 (21.2)	80.0 (20.3)	0.96
Systolic BP (mmHg)	132.5 (18.5)	128.4 (20.8)	<0.0001	132.3 (18.6)	135.7 (18.5)	0.01	132.4 (18.6)	130.9 (20.0)	0.12
Diastolic BP (mmHg)	78.2 (11.9)	74.6 (12.9)	<0.0001	78.1 (11.9)	77.3 (12.4)	0.33	78.1 (11.9)	75.7 (12.4)	<0.0001
Scores									
CHA_2DS_2-VASc	3.15 (1.5)	4.1 (1.6)	<0.0001	3.18 (1.5)	3.95 (1.7)	<0.0001	3.17 (1.5)	3.83 (1.5)	<0.0001
HAS-BLED	1.32 (0.9)	1.7 (0.9)	<0.0001	1.33 (0.9)	1.68 (0.9)	<0.0001	1.33 (0.9)	1.69 (0.9)	<0.0001
CCI*	3.84 (1.9)	4.53 (1.8)	<0.0001	3.84 (1.9)	4.53 (1.8)	<0.0001	3.83 (1.9)	4.86 (2.2)	<0.0001
Laboratory result									
Creatinine clearance (mL/min)	82.2 (52.7)	62.9 (32.6)	<0.0001	81.6 (52.4)	71.2 (32.7)	0.004	81.7 (52.5)	70.7 (30.5)	<0.0001
Comorbidities									
Hypertension, No. (%)	18 653 (75.3)	673 (75.0)	0.83	19 149 (75.3)	177 (82.3)	0.02	19 000 (75.2)	326 (80.5)	0.02
Hypercholesterolemia, No. (%)	9951 (40.2)	345 (38.5)	0.3	10 198 (40.1)	98 (45.6)	0.1	10 088 (40)	208 (51.4)	<0.0001
Diabetes mellitus, No. (%)	5670 (22.9)	266 (29.7)	<0.0001	5882 (23.1)	54 (25.1)	0.49	5822 (23.1)	114 (28.1)	0.02
Coronary artery disease, No. (%)	4628 (18.7)	269 (30)	<0.0001	4842 (19)	55 (25.6)	0.01	4792 (19)	105 (25.9)	0.0004
Congestive heart failure, No. (%)	5304 (21.4)	396 (44.1)	<0.0001	5653 (22.2)	47 (21.9)	0.9	5583 (22.1)	117 (28.9)	0.001
Left ventricular hypertrophy, No. (%)	4528 (18.3)	197 (22)	0.01	4683 (18.4)	42 (19.5)	0.67	4648 (18.4)	77 (19)	0.76
Prior thromboembolism, No. (%)	3615 (14.6)	205 (22.9)	<0.0001	3741 (14.7)	79 (36.7)	<0.0001	3739 (14.8)	81 (20)	0.004
Prior bleeding, No. (%)	1285 (5.2)	59 (6.6)	0.07	1335 (5.2)	9 (4.2)	0.49	1301 (5.2)	43 (10.6)	<0.0001
Peripheral artery disease	686 (2.8)	61 (6.8)	<0.0001	737 (2.9)	10 (4.7)	0.13	726 (2.9)	21 (5.2)	0.01
COPD, No. (%)	1409 (5.7)	124 (13.8)	<0.0001	1518 (6)	15 (7)	0.53	1485 (5.9)	48 (11.9)	<0.0001
Dementia	138 (0.6)	6 (0.7)	0.0004	142 (0.6)	2 (0.9)	0.02	139 (0.6)	5 (1.2)	<0.0001
Peripheral vascular disease	3138 (12.7)	73 (8.2)	<0.0001	3199 (12.6)	12 (5.6)	0.39	3177 (12.6)	34 (8.4)	<0.0001
Cancer	2424 (9.8)	42 (4.7)	0.002	2459 (9.7)	7 (3.3)	0.90	2447 (9.7)	19 (4.7)	0.13
Metastatic cancer	116 (0.5)	6 (0.7)	0.005	122 (0.5)	0 (0.0)	0.55	119 (0.5)	3 (0.7)	0.03
Peptic ulcer disease	419 (1.7)	5 (0.6)	0.85	422 (1.7)	2 (0.9)	0.50	422 (1.7)	2 (0.5)	0.83
Liver disease	365 (1.5)	11 (1.2)	0.0006	376 (1.5)	0 (0.0)	0.29	374 (1.5)	2 (0.5)	0.96
Chronic kidney disease	370 (1.5)	34 (3.8)	<0.0001	400 (1.6)	4 (1.9)	0.01	398 (1.6)	6 (1.5)	0.01
Treatment									
OAC, No. (%)	21 266 (85.9)	711 (79.3)	<0.0001	21 812 (85.7)	165 (76.7)	0.0002	21 614 (85.6)	363 (89.6)	0.02
Dabigatran	8326 (33.6)	213 (23.7)		8477 (33.3)	62 (28.8)		8428 (39.0)	111 (27.4)	
VKA	4547 (18.4)	226 (25.2)		4737 (18.6)	36 (16.7)		4656 (21.5)	117 (28.9)	

Continued

Table 1 *Continued*

	All-cause death			Ischaemic stroke			Major bleeding		
	No event (*n* = 24 759)	Event (*n* = 897)	*P*-value	No event (*n* = 25 441)	Event (*n* = 215)	*P*-value	No event (*n* = 25 251)	Event (*n* = 405)	*P*-value
Apixaban	4274 (17.3)	134 (14.9)		4368 (17.2)	40 (18.6)		4347 (20.1)	61 (15.1)	
Rivaroxaban	3803 (15.4)	130 (14.5)		3908 (15.4)	25 (11.6)		3863 (17.9)	70 (17.3)	
Edoxaban	316 (1.3)	8 (0.9)		322 (1.3)	2 (0.9)		320 (1.5)	4 (1.0)	
AF ablation, No. (%)	438 (1.8)	4 (0.4)	0.003	441 (1.7)	1 (0.5)	0.15	439 (1.7)	3 (0.7)	0.13
Antiplatelet, No. (%)	5923 (23.9)	276 (30.8)	<0.0001	6139 (24.1)	60 (27.9)	0.2	6070 (24)	129 (31.9)	0.0003
Antiarrhythmic drug, No. (%)	6399 (25.8)	226 (25.2)	0.66	6592 (25.9)	33 (15.3)	0.0004	6530 (25.9)	95 (23.5)	0.27
ACE-i, No. (%)	7534 (30.4)	289 (32.2)	0.25	7765 (30.5)	58 (27)	0.26	7689 (30.5)	134 (33.1)	0.25
Angiotensin receptor blocker, No. (%)	6393 (25.8)	214 (23.9)	0.19	6554 (25.8)	53 (24.7)	0.71	6498 (25.7)	109 (26.9)	0.59
Beta blocker, No. (%)	15 579 (62.9)	580 (64.7)	0.29	16 025 (63)	134 (62.3)	0.84	15 881 (62.9)	278 (68.6)	0.02
Digoxin, No. (%)	2001 (8.1)	119 (13.3)	<0.0001	2095 (8.2)	25 (11.6)	0.07	2075 (8.2)	45 (11.1)	0.04
Diuretic, No. (%)	9255 (37.4)	493 (55)	<0.0001	9673 (38)	75 (34.9)	0.35	9555 (37.8)	193 (47.7)	0.0001
Statin, No. (%)	11 104 (44.8)	430 (47.9)	0.07	11 406 (44.8)	128 (59.5)	<0.0001	11 311 (44.8)	223 (55.1)	<0.0001
AF classification									
Type of AF, No. (%)									
Paroxysmal	13 807 (55.8)	429 (47.8)	<0.0001	14 135 (55.6)	101 (47.0)	0.032	14 027 (55.6)	209 (51.6)	0.284
Persistent	8482 (34.3)	339 (37.8)		8730 (34.3)	91 (42.3)		8670 (34.3)	151 (37.3)	
Permanent	2470 (10.0)	129 (14.4)		2576 (10.1)	23 (10.7)		2554 (10.1)	45 (11.1)	
EHRA classification, No. (%)									
I	8399 (33.9)	350 (39.0)	<0.0001	8653 (34.0)	96 (44.7)	0.007	8586 (34.0)	163 (40.3)	0.003
II	10 085 (40.7)	278 (31.0)		10 297 (40.5)	66 (30.7)		10 225 (40.5)	138 (34.1)	
III	4853 (19.6)	215 (24.0)		5027 (19.8)	41 (19.1)		4997 (19.8)	71 (17.5)	
IV	1422 (5.7)	54 (6.0)		1464 (5.8)	12 (5.6)		1443 (5.7)	33 (8.2)	

*CCI was derived using characteristics including age, myocardial infarction, congestive heart feature, peripheral vascular disease, cerebrovascular accident or transient ischaemic attack, dementia, COPD, peptic ulcer disease, liver disease, diabetes mellitus, chronic kidney disease, cancer, and metastatic cancer. Information for connective tissue, hemiplegia, leukaemia, lymphoma, and AIDS were not available in the data.

ACE-i, angiotensin-converting enzyme inhibitor; AF, atrial fibrillation; BMI, body mass index; BP, blood pressure; CCI, Charlson Comorbidity Index; COPD, chronic obstructive pulmonary disease; EHRA, European Heart Rhythm Association; OAC, oral anticoagulant; SD, standard deviation; VKA, vitamin K antagonist; AIDS, acquired immunodeficiency syndrome.

Results

The study cohort comprised 25 656 patients in total, which 10 961 patients were excluded due to a lack of follow-up information. The patients were predominantly anticoagulated (86.0%) with a mean age of 70.3 years (SD 10.3), and 44.8% female, and with a mean CHA_2DS_2-VASc score of 3.2 (SD 1.5) (see Supplementary material online, *Table S3* for characteristics of the total cohort, and Supplementary material online, *Tables S4–S7* for characteristics of the excluded patients and sub-cohorts).

During the first-year follow-up, there were 215 (0.8%) ischaemic stroke events with a median time-to-event 139 days, 897 (3.5%) all-cause death with a median time-to-event 158 days, and 405 (1.6%) major bleeding events with median time 138 days. Among patients who received anticoagulant therapy, the incidence of ischaemic stroke was 0.75 per 100 patient-years (PYs), the incidence of all-cause death was 3.2 per 100 PYs, and the incidence of major bleeding was 1.7 per 100 PYs.

The incidence of all-cause death varied across different regions, with incidence of 3.7 per 100 PYs among patients residing in Europe, 3.5 per 100 PYs for patients in North America, 2.3 per 100 PYs for patients residing in Asia, and 5.8 per 100 PYs for patients in Latin America. *Table 1* presents a comparison of the baseline characteristics of patients with ischaemic stroke with those who had major bleeding or died during the same period. The use of oral anticoagulants (OACs) was lower in patients who suffered an ischaemic stroke (76.7 vs. 85.7%) or died (79.3 vs. 85.9%), but higher in patients with major bleeding (89.6 vs. 85.6%).

Performance

The predictive ability of the model for all-cause death was significantly better for ML-GBDT compared with CCI, with an AUC-ROC of 0.785 (95% CI: 0.757–0.813) vs. 0.747 (0.718–0.777), respectively ($P = 0.007$) (*Figure 1*). We also evaluated the performance of CHA_2DS_2-VASc in predicting all-cause death, and it performed worse than the CCI (0.676 vs. 0.747, $P < 0.001$). The NRI for predicting all-cause death was 10.0% on the total cohort using our ML-GBDT model. Correspondingly, the decision curve for ML-GBDT showed a higher net benefit than the CCI across the majority of the threshold probability range. Relative to CCI, the ML-GBDT yielded a higher sensitivity (0.726 vs. 0.677) while maintaining comparable specificity (0.707 vs. 0.704).

The AUC-ROC for predicting ischaemic stroke was significantly higher ($P = 0.028$) for ML-GBDT [0.691 (0.626–0.756)] compared with CHA_2DS_2-VASc [0.613 (0.538–0.689)]. The NRI showed an

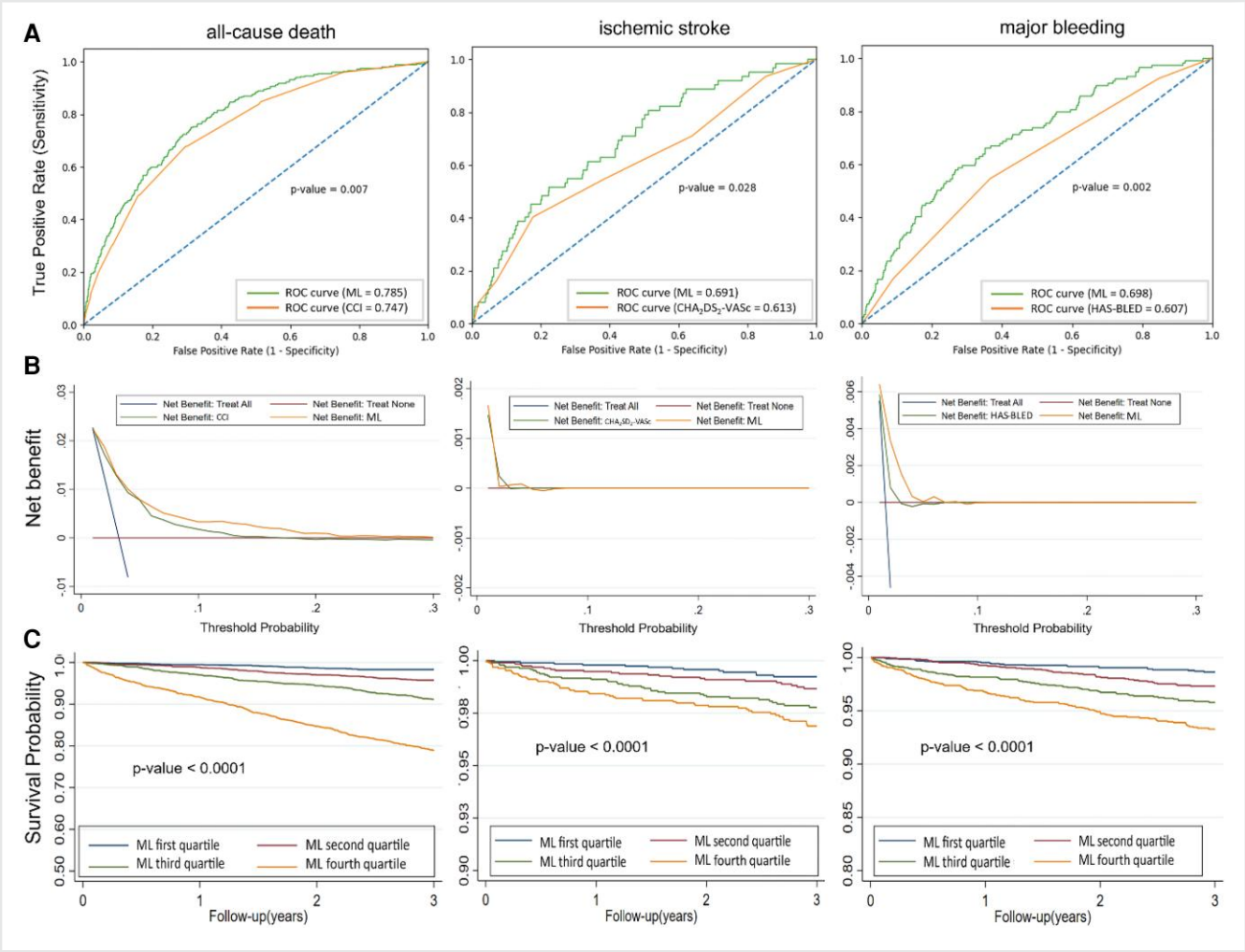


Figure 1 Performance of machine learning model compared with that of clinical scores in predicting risk of all-cause death, ischaemic stroke, and major bleeding in patients with atrial fibrillation. (*A*) Area under the receiver operating characteristic curve and (*B*) decision analysis curve. The *x*-axis represents the threshold probability, and the *y*-axis illustrates the net benefit derived from each model. The 'treat all' and 'treat none' strategies are also depicted on the decision curve. (*C*) Kaplan–Meier survival curves. The survival curves plot time in years on the *x*-axis, against the survival probability on the *y*-axis. The predicted probabilities were grouped into quartiles, with each represented by a distinct survival curve, corresponding to the respective predicted probability range. ROC, receiver operating characteristic; ML, machine learning.

improvement of 12.5%. A similar specificity for both the model and score were obtained while choosing the probability thresholds with ML-GBDT consistently demonstrating superior sensitivity.

An improvement was also observed in predicting major bleeding, with the AUC-ROC significantly ($P = 0.002$) higher for ML-GBDT than HAS-BLED: 0.698 (0.651–0.745) compared with 0.607 (0.559–0.655). The ML-GBDT model yielded a higher sensitivity (0.664 vs. 0.546) when predicting major bleeding. Of the three outcomes, the NRI for major bleeding was the highest at 23.6%, in alignment with the decision curve analysis (*Figure 1*). The survival curve (*Figure 1*) for the 'low risk' (first and second quartiles) groups were observed to have a slower decline, indicating higher survival rates over time, compared with the 'high risk' (third and fourth quartiles) groups. It also outperformed single-label GBDT and Cox regression model in predicting the risks for patients independently (see Supplementary material online, *Table S8*). Although we limited the number of variables to 25 in our model for clinical application, it maintained a similar AUC in predicting all-cause death [0.787 (0.763, 0.811)], ischaemic stroke [0.711 (0.654, 0.768)], and major bleeding [0.686 (0.637, 0.734)].

As shown in *Table 2*, the performance of ML-GBDT predicting patients' risk within 90 days mirrored its performance in the 1-year follow-up. CCI also achieved similar performance in predicting short-term all-cause mortality risks for patients. However, AUC-ROC of CHA_2DS_2-VASc and HAS-BLED decreased when predicting short-term risks compared with their long-term prediction.

Sub-group analyses

The ML-GBDT model demonstrated superior discriminatory capacity than the CCI ($P < 0.05$) on all sub-cohorts, excluding patients for secondary stroke prevention. The CCI showed a reduced AUC-ROC while predicting all-cause death in patients who were older (age ≥ 65), anticoagulated, and residing in European, North American, or Asian regions, compared with the total cohort (*Table 3* and *Figure 2*). For patients who were not anticoagulated, those from Europe, or patients for primary or secondary stroke prevention, AUC-ROC of CCI was comparable or higher than the total cohort.

Table 2 Evaluation of the performance of multi-label gradient boosting decision tree and clinical risk scores at different time point during follow-up in the total cohort

		CCI (95% CI)	CHA_2DS_2-VASc (95% CI)	HAS-BLED (95% CI)	ML-GBDT (95% CI)
90 days					
All-cause death ($n = 279$)	AUC	0.747 (0.718, 0.777)	0.498 (0.439, 0.556)		0.786 (0.743, 0.830)
	SENS	0.661 (0.611, 0.724)			0.727 (0.657, 0.837)
	SPEC	0.697 (0.687, 0.708)			0.707 (0.697, 0.718)
Ischaemic stroke ($n = 98$)	AUC		0.476 (0.378, 0.575)		0.698 (0.589, 0.807)
	SENS		0.500 (0.374, 0.754)		0.680 (0.558, 0.955)
	SPEC		0.605 (0.594, 0.616)		0.618 (0.608, 0.629)
Major bleeding ($n = 76$)	AUC			0.481 (0.403, 0.558)	0.633 (0.550, 0.716)
	SENS			0.325 (0.220, 0.526)	0.558 (0.448, 0.758)
	SPEC			0.618 (0.607, 0.629)	0.632 (0.622, 0.643)
1 year					
All-cause death ($n = 897$)	AUC	0.747 (0.718, 0.777)	0.676 (0.643, 0.709)		0.785 (0.757, 0.813)
	SENS	0.677 (0.617, 0.733)			0.726 (0.667, 0.778)
	SPEC	0.704 (0.693, 0.714)			0.707 (0.697, 0.717)
Ischaemic stroke ($n = 215$)	AUC		0.613 (0.538, 0.689)		0.691 (0.626, 0.756)
	SENS		0.548 (0.425, 0.666)		0.629 (0.505, 0.738)
	SPEC		0.614 (0.603, 0.625)		0.617 (0.606, 0.628)
Major bleeding ($n = 405$)	AUC			0.607 (0.559, 0.655)	0.698 (0.651, 0.745)
	SENS			0.546 (0.457, 0.633)	0.664 (0.575, 0.742)
	SPEC			0.634 (0.623, 0.644)	0.634 (0.623, 0.645)

The AUC-ROC of the CHA_2DS_2-VASc score was higher among elderly, anticoagulated, European and North American sub-cohorts, as opposed to its AUC-ROC for the total cohort while estimating the risks of ischaemic stroke (*Table 3*). However, its performance declined in patients from Asia or those for either primary or secondary stroke prevention, relative to the total cohort. The ML-GBDT model outperformed the CHA_2DS_2-VASc score in predicting ischaemic stroke across all the sub-groups, but none were statistically significant.

The AUC-ROC of HAS-BLED remained consistent or increased among several sub-groups of patients when compared with its AUC-ROC for the total cohort. These sub-groups included those who received anticoagulant therapy ($n = 21\,987$), those who were from European ($n = 12\,792$) and North American regions ($n = 5974$), and patients for primary stroke prevention ($n = 21\,912$). Conversely, it was lower for elderly patients, those not receiving anticoagulation therapy ($n = 3678$), patients from Asia ($n = 4769$), and those for secondary stroke prevention ($n = 3744$), in comparison to the total cohort. The ML-GBDT achieved higher AUC-ROC than the HAS-BLED, but not all observed predictions were statistically different.

Figure 3 and *Figure 4* illustrate the top 25 or 30 most important variables for total and sub-cohorts in the ML-GBDT model. Age, smoking, creatinine clearance, anticoagulant therapy, and history of stroke or transient ischaemic attack (TIA) emerged as the most critical variables in predicting the risks for patients diagnosed with AF. In the sub-cohorts, the ranking of top features changed, including blood pressure, regions, body mass index (BMI), previous thromboembolism, diabetes, respiratory disease, congestive heart failure, and statin therapy. These additional variables, along with age, creatinine clearance, and history of stroke/TIA, significantly impacted risk prediction. Furthermore, we have transformed the categorical variable 'OAC type' with one-hot encoding and evaluated their feature importance in the retrained model. Vitamin K antagonist emerged as the most influential OAC treatment on patients' risks, followed by dabigatran, while apixaban, rivaroxaban, and edoxaban were deemed less important. A similar approach was applied to analyse the importance of different regions and races. Detailed feature ranking is listed in Supplementary material online, *Figures S1–3*.

Discussion

In this study, we have shown that our ML-GBDT model performed better than three clinical risk scores in predicting both short- (90-day) and long-term (1-year) risk of death, ischaemic stroke, and major bleeding, for patients diagnosed with AF. This attained statistically significant optimized performance in comparison to the CCI for predicting the risk of all-cause death, the CHA_2DS_2-VASc score for ischaemic stroke risk, and the HAS-BLED score for major bleeding risk. Moreover, our study is the first to present the feature importance ranking for patient cohorts featured by geographical regions, receiving anticoagulant therapy, and history of stroke.

During the past decades, there has been growing recognition that death, especially cardiovascular death, associated with AF is increasing,[2] although there is no dedicated score for this specific population. It is therefore important to implement appropriate measures to mitigate the risk of death for patients with AF. Our developed ML-GBDT model provides statistically significant improved performance in assessing the risk of all-cause death accounting for the risk of ischaemic stroke and major bleeding.

Whilst OAC therapy is used to reduce the risk of stroke, it simultaneously increases the likelihood of major bleeding events.[3,26] Patients may still experience an ischaemic stroke event despite receiving anticoagulant therapy.[5] Also, despite an uptake in the use of anticoagulants, one-third of high-risk patients with AF do not receive any anticoagulant

Table 3 Comparison of the area under the receiver operating characteristic curve of the multi-label gradient boosting decision tree model and clinical risk scores across various study sub-groups

	CCI (95% CI)	CHA_2DS_2-VASc (95% CI)	HAS-BLED (95% CI)	ML-GBDT (95% CI)	*P*-value (score vs. ML model)
Age ≥ 65 years old ($n = 19\,032$)					
All-cause death ($n = 792$)	0.712 (0.679, 0.746)			0.752 (0.718, 0.785)	0.0283
Ischaemic stroke ($n = 184$)		0.653 (0.575, 0.730)		0.715 (0.652, 0.778)	0.0913
Major bleeding ($n = 352$)			0.548 (0.492, 0.605)	0.627 (0.572, 0.682)	0.0214
Anticoagulated ($n = 21\,987$)					
All-cause death ($n = 711$)	0.720 (0.687, 0.752)			0.759 (0.728, 0.790)	0.0027
Ischaemic stroke ($n = 165$)		0.698 (0.622, 0.774)		0.750 (0.677, 0.823)	0.0772
Major bleeding ($n = 363$)			0.601 (0.554, 0.647)	0.685 (0.637, 0.733)	0.0018
Not anticoagulated ($n = 3678$)					
All-cause death ($n = 186$)	0.759 (0.703, 0.816)			0.819 (0.766, 0.872)	0.0396
Ischaemic stroke ($n = 50$)		0.620 (0.475, 0.765)		0.694 (0.593, 0.796)	0.2394
Major bleeding ($n = 42$)			0.584 (0.416, 0.752)	0.678 (0.540, 0.815)	0.2838
Europe ($n = 12\,792$)					
All-cause death ($n = 467$)	0.747 (0.711, 0.783)			0.796 (0.763, 0.828)	0.0014
Ischaemic stroke ($n = 108$)		0.681 (0.600, 0.762)		0.737 (0.663, 0.810)	0.2102
Major bleeding ($n = 199$)			0.645 (0.581, 0.709)	0.690 (0.611, 0.753)	0.3232
North America ($n = 5974$)					
All-cause death ($n = 201$)	0.723 (0.663, 0.783)			0.804 (0.742, 0.865)	0.0050
Ischaemic stroke ($n = 40$)		0.696 (0.515, 0.878)		0.705 (0.530, 0.880)	0.8928
Major bleeding ($n = 143$)			0.653 (0.564, 0.742)	0.708 (0.626, 0.789)	0.3186
Asia ($n = 4769$)					
All-cause death ($n = 108$)	0.692 (0.584, 0.799)			0.840 (0.766, 0.914)	0.0123
Ischaemic stroke ($n = 54$)		0.532 (0.390, 0.675)		0.676 (0.515, 0.836)	0.1419
Major bleeding ($n = 43$)			0.582 (0.452, 0.713)	0.606 (0.467, 0.746)	0.7973
Primary prevention ($n = 21\,912$)					
All-cause death ($n = 698$)	0.744 (0.714, 0.773)			0.782 (0.750, 0.814)	0.0104
Ischaemic stroke ($n = 137$)		0.527 (0.455, 0.598)		0.598 (0.518, 0.677)	0.1343
Major bleeding ($n = 327$)			0.623 (0.569, 0.676)	0.656 (0.604, 0.709)	0.2380
Secondary prevention ($n = 3744$)					
All-cause death ($n = 199$)	0.743 (0.681, 0.805)			0.773 (0.707, 0.839)	0.4192
Ischaemic stroke ($n = 78$)		0.532 (0.440, 0.624)		0.627 (0.530, 0.724)	0.2319
Major bleeding ($n = 78$)			0.556 (0.414, 0.697)	0.630 (0.518, 0.743)	0.3848

The ML-GBDT was trained and tested on the sub-cohort.
AUC-ROC, area under the receiver operating characteristic curve; CI, confidence interval; CCI, Charlson Comorbidity Index; ML-GBDT, multi-label gradient boosting decision tree.

treatment.[27] As a result, the risk of death, stroke, and major bleeding should be monitored to identify patients at increased risk so that appropriate intervention or closer monitoring can be initiated. The low number of events, specifically ischaemic stroke, resulting from the use of anticoagulant therapy, is a challenge for prediction. This has indeed impacted the predictive performance for both clinical scores and our model.

Using clinical risk scores, anticoagulant therapy is recommended in patients with a CHA_2DS_2-VASc score ≥ 2 in males or ≥3 for females and should be considered when CHA_2DS_2-VASc score is 1 in males or 2 for females. Meanwhile, patients at high risk of bleeding (HAS-BLED score ≥ 3) should be scheduled for early and more frequent clinical review and follow-up.[9] However, scores which only include a limited number of risk factors report a modest capacity in predicting patient risks, particularly in predicting short-term risks of ischaemic stroke and major bleeding. Our model which additionally incorporated treatments, interventions, and an extensive range of comorbidities, yielded the best results. It also demonstrated its reliability and robustness, as evidenced by its consistent performance in predicting both short- and long-term risks for patients with AF. By leveraging the predicted risk probability and feature importance, clinicians could address modifiable risk factors and adjust treatments and interventions utilizing our model. Fox *et al.*[28] have built a GARFIELD-AF risk tool for predicting all-cause mortality, non-haemorrhagic stroke, and major bleeding. It also outperformed CHA_2DS_2-VASc and HAS-BLED, which is consistent with our study. While both studies targeted newly diagnosed AF patients with at least one stroke risk factors, our model demonstrated higher AUC-ROC in predicting all-cause mortality (GARFIELD-AF 0.75 vs. 0.79 ML-GBDT), particularly among patients not receiving anticoagulant treatment (GARFIELD-AF 0.77 vs. 0.82 ML-GBDT). However, the differing cohort size and follow-up durations between our study and GARFIELD-AF must be considered while interpreting the performance.

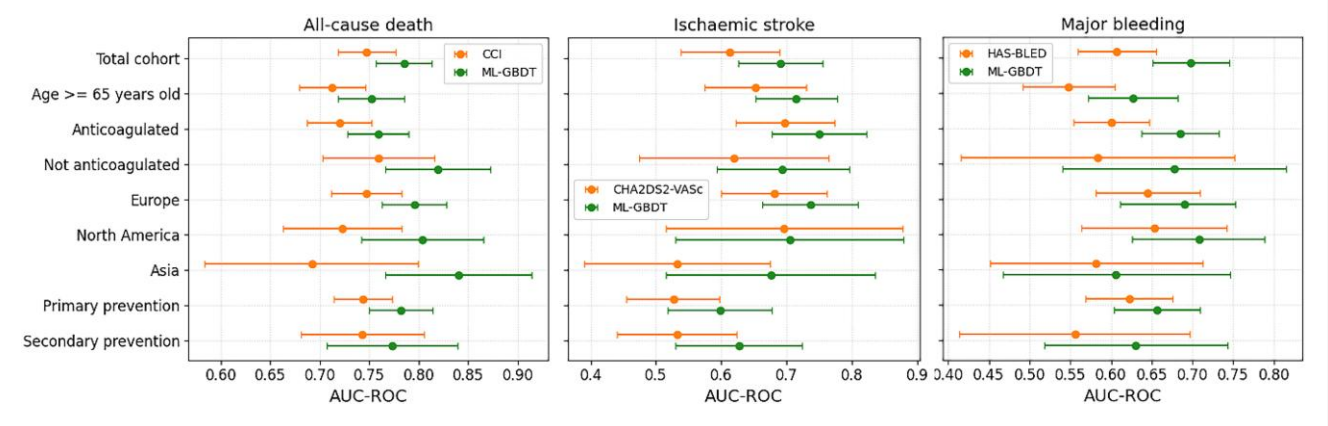


Figure 2 Comparison of area under the receiver operating characteristic curve for clinical scores and multi-label gradient boosting decision tree in predicting risk of all-cause death, ischaemic stroke, and major bleeding. It illustrates area under the receiver operating characteristic curve values along with their respective confidence intervals, demonstrating the performance of the total cohort as well as specific sub-groups.

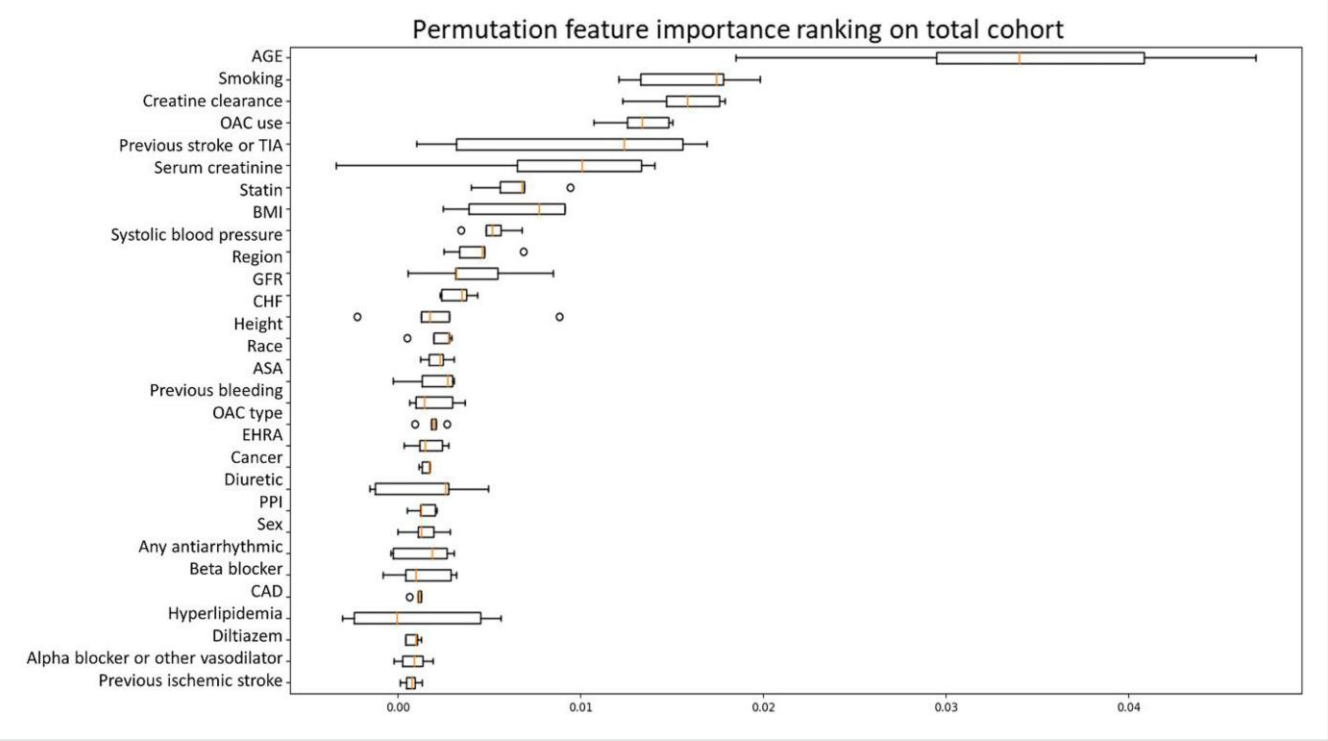


Figure 3 Permutation feature importance in the multi-label gradient boosting decision tree model on the total cohort while predicting outcomes of all-cause death, ischaemic stroke, and major bleeding. ASA, aspirin; BMI, body mass index; CAD, coronary artery disease; CHF, congestive heart failure; EHRA, European Heart Rhythm Association; GFR, glomerular filtration rate; OAC, oral anticoagulant; PPI, proton pump inhibitor; TIA, transient ischaemic attack.

AF necessitates a multifaceted, holistic approach to its management and requires active patient involvement in partnership with clinicians. The Atrial Fibrillation Better Care pathway was introduced as an approach to simplify the integrated care in patients with AF,[29] and adhering to it has been shown to improve the prognosis of AF patients.[30,31] Current guidelines stipulate that a holistic or integrated care approach to AF management should identify and manage risk factors and comorbidities.[9,10]

With insights from feature importance in the ML-GBDT model, it is additionally possible to characterize subclinical processes and target key modifiable risk factors to modify patient risks, such as hypertension, diabetes, and BMI. This could potentially enrich the precision of interventions and lead to optimized patient care. The ML-GBDT model and its feature importance could further be integrated with electronic medical records or mobile apps in the future in daily clinical practice to facilitate

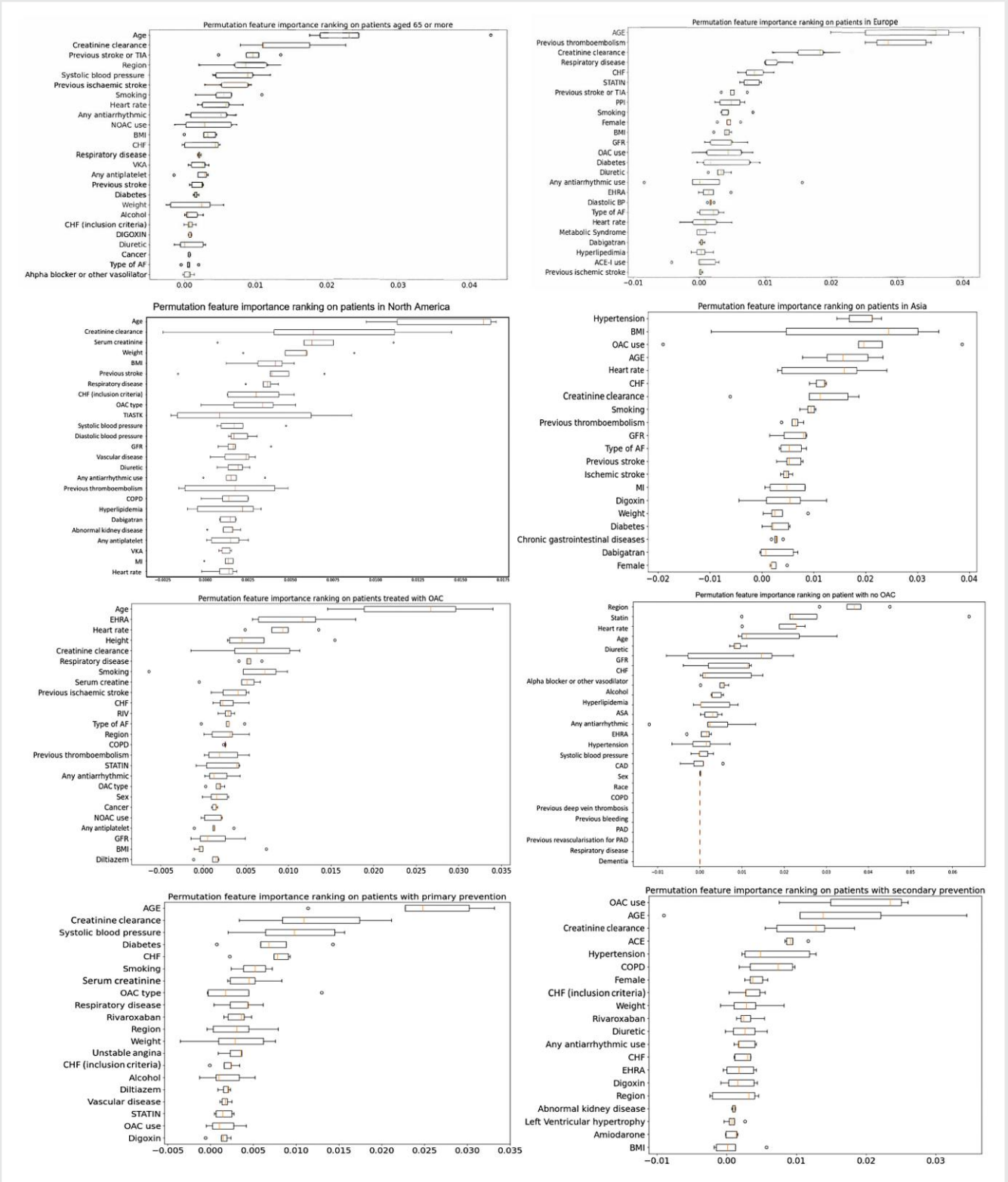


Figure 4 Permutation feature importance rankings in the multi-label gradient boosting decision tree model on the sub-cohort while predicting outcomes of all-cause death, ischaemic stroke, and major bleeding. ACE-i, angiotensin-converting enzyme inhibitor; AF, atrial fibrillation; ASA, aspirin; BMI, body mass index; CAD, coronary artery disease; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; EHRA, European Heart Rhythm Association; GFR, glomerular filtration rate; MI, myocardial infarction; NOAC, non-vitamin K antagonist oral anticoagulant; OAC, oral anticoagulant; PAD, peripheral artery disease; PPI, proton pump inhibitor; RIV, rivaroxaban; STK, stroke; TIA, transient ischaemic attack; VKA, vitamin K antagonist.

better stratification of risk and improve the prognosis of patients through tailored management.

Anticoagulation

Despite receiving anticoagulation treatment, patients with AF may still face an enduring risk of stroke and mortality. Residual risks for death and ischaemic stroke were predicted by our model in patients who received anticoagulant therapy. As indicated in *Table 2*, the use of anticoagulant therapy had a significant impact on the performance of AF patient risk prediction.

Most patients in the GLORIA-AF registry received anticoagulant therapy due to their elevated risk of ischaemic stroke (CHA_2DS_2-VASc score ≥ 1), and the number of patients who were not anticoagulated was small, resulting in a less statistically significant finding while comparing the performance of model and scores in the latter group. Notably, more than three-quarters of anticoagulated patients were under NOAC in our cohort, and about 50% of these patients received Dabigatran, known for its lower risk of major bleeding.[32] The prevalent use of NOAC may have contributed to the lower observed incidence of major bleeding. Consequently, this reduction could have impacted HAS-BLED in accurately predicting patients' risks for major bleeding.

Patient location

Our results indicated that the location where patients reside was statistically significant for each outcome ($P < 0.0001$) and influenced the risk prediction performance. The AUC-ROCs for all the clinical risk scores were lower for patients from Asia. This could potentially be explained by the age difference across regions (see Supplementary material online, *Table S4*). Chao *et al.*[14] showed that a modified CHA_2DS_2-VASc optimized performance for stroke risk prediction in Asian patients, which reflected the age threshold in Asian patients with AF for an increased risk of ischaemic stroke could be 50 years.[33]

Primary vs. secondary prevention sub-groups

Stroke history is a critical risk factor while assessing a patients' risk of ischaemic stroke. Both ML models and clinical risk scores yielded suboptimized performance in predicting the risks for patients who have not previously had a stroke. Lower event rates were observed in these patients (0.63%) compared with 2.1% for secondary prevention (see Supplementary material online, *Table S5*). This was lower than the observed event rate in the Darlington AF registry,[34] but this could be due to more than 85% of patients in the GLORIA-AF registry receiving anticoagulant therapy. The OAC use was also ranked as one of the top important factors for patients with secondary prevention of stroke in *Figure 3*.

Machine learning

Machine learning has been increasingly applied to predict risks for patients with AF using various data sources.[35–38] Ambale-Venkatesh *et al.* employed different data sources, including imaging, text, questionnaires, and biomarkers, to improve cardiovascular event predictions through ML algorithms.[36] ML has also been used to uncover risk factors that were previously unknown and improve stroke prediction in patients with AF.[37] Sequential measures of international normalized ratio have also been used to predict clinical outcomes in newly diagnosed patients using ML.[38] Compared with other applications of ML, our ML-GBDT model took into account the relationship between different outcomes (all-cause death, ischaemic stroke, and major bleeding) for patients with AF. We have also compared the feature importance ranking of different subcohorts, which could provide additional support in practice.

Application

Our approach to clinical implementation focused on making risk assessment more accessible and accurate for healthcare practitioners. The link to our online risk calculator is available in Supplementary material online. By limiting the variables to 25 for clinical application, we maintained robust model performance without overwhelming users. This strategy emphasizes our tool balances comprehensiveness and ease of use, ensuring it can be efficiently integrated into clinical workflows.

Limitations

A key limitation in our study cohort is the relatively low number of events, especially ischaemic stroke, due to the use of anticoagulant therapy. The current study does not account for possible changes in anticoagulant status during follow-up, nor the impact of clinical interventions that impact death, stroke, or bleeding, such as early rhythm control.[39,40] Another limitation to the predictive performance could be the structure of the classifier chains. The sequential nature of the model, where each subsequent classifier's input includes the output of the preceding one, implies that information propagates along the chain, including both performance gains and inaccuracies. Although inaccuracies may affect the prediction, the information propagation may have ultimately led to enhanced performance. One of the challenges associated with the application of ML models is the requirement for a much larger number of variables compared with risk scores because it accounts for an extensive range of patient comorbidities, current treatments, and interventions. This might pose an additional workload to collect these variables in already busy clinical settings. However, the ML process has the capability of full automation, by integrating the model directly into electronic health record systems. The whole process, from data extraction to risk prediction could occur in real-time. This could substantially mitigate the associated burden of the required data collection, making the adoption of models more feasible and less disruptive in future.

Conclusions

Our study validates that the multi-label ML model outperforms the clinical risk scores in predicting the risk of all-cause death, ischaemic stroke, and major bleeding in patients with AF from a contemporary global registry. We also provide feature importance ranking for various patient cohorts by geographical regions, receiving anticoagulation therapy, and history of stroke. Our approach could be used as a multifaceted and holistic tool to optimize patient risk assessments and mitigate adverse outcomes in diverse AF populations.

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health*.

Conflict of interest: none declared.

Data availability

The data underlying this article will be shared on reasonable request to the GLORIA-AF chairs, Prof. Lip and Prof. Huisman.

References

1. Kornej J, Börschel CS, Benjamin EJ, Schnabel RB. Epidemiology of atrial fibrillation in the 21st century. *Circ Res* 2020;**127**:4–20.
2. Tanaka Y, Shah NS, Passman R, Greenland P, Lloyd-Jones DM, Khan SS. Trends in cardiovascular mortality related to atrial fibrillation in the United States, 2011 to 2018. *J Am Heart Assoc* 2021;**10**:e020163.
3. Ruff CT, Giugliano RP, Braunwald E, Hoffman EB, Deenadayalu N, Ezekowitz MD, *et al.* Comparison of the efficacy and safety of new oral anticoagulants with warfarin in

patients with atrial fibrillation: a meta-analysis of randomised trials. *Lancet* 2014;383: 955–962.
4. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007;**146**:857–867.
5. Ding WY. Residual stroke risk in atrial fibrillation. *Arrhythmia Electrophysiol Rev* 2021;**10**: 147–153.
6. Ding WY, Lane DA, Gupta D, Huisman MV, Lip GYH, on behalf of the G-AI. Incidence and risk factors for residual adverse events despite anticoagulation in atrial fibrillation: results from phase II/III of the GLORIA-AF registry. *J Am Heart Assoc* 2022;**11**:e026410.
7. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the Euro heart survey on atrial fibrillation. *Chest* 2010;**137**: 263–272.
8. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJGM, Lip GYH. 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**:1093–1100.
9. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, *et al.* 2020 ESC guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). *Eur Heart J* 2020;**42**:ehaa612.
10. Chao T-F, Joung B, Takahashi Y, Lim TW, Choi E-K, Chan Y-H, *et al.* 2021 focused update consensus guidelines of the Asia Pacific Heart Rhythm Society on stroke prevention in atrial fibrillation: executive summary. *Thromb Haemost* 2022;**122**:020–047.
11. Chen J-Y, Zhang A-D, Lu H-Y, Guo J, Wang F-F, Li Z-C. $CHADS_2$ 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. *J Geriatr Cardiol* 2013;**10**:258–266.
12. Senoo K, Proietti M, Lane DA, Lip GYH. Evaluation of the HAS-BLED, ATRIA, and ORBIT bleeding risk scores in patients with atrial fibrillation taking warfarin. *Am J Med* 2016;**129**:600–607.
13. van Doorn S, Debray TPA, Kaasenbrood F, Hoes AW, Rutten FH, Moons KGM, *et al.* Predictive performance of the CHA_2DS_2-VASc rule in atrial fibrillation: a systematic review and meta-analysis. *J Thromb Haemost* 2017;**15**:1065–1077.
14. Chao T-F, Lip GYH, Liu C-J, Tuan T-C, Chen S-J, Wang K-L, *et al.* Validation of a modified CHA_2DS_2-VASc score for stroke risk stratification in Asian patients with atrial fibrillation. *Stroke* 2016;**47**:2462–2469.
15. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; **40**:373–383.
16. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson comorbidity index: a critical review of clinimetric properties. *Psychother Psychosom* 2022;**91**:8–35.
17. Lip GYH, Genaidy A, Tran G, Marroquin P, Estes C, Sloop S. Improving stroke risk prediction in the general population: a comparative assessment of common clinical rules, a new multimorbid index, and machine-learning-based algorithms. *Thromb Haemost* 2022; **122**:142–150.
18. Lu J, Hutchens R, Hung J, Bennamoun M, McQuillan B, Briffa T, *et al.* Performance of multilabel machine learning models and risk stratification schemas for predicting stroke and bleeding risk in patients with non-valvular atrial fibrillation. *Comput Biol Med* 2022; **150**:106126.
19. Huisman MV, Lip GYH, Diener HC, Dubner SJ, Halperin JL, Ma CS, *et al.* Design and rationale of global registry on long-term oral antithrombotic treatment in patients with atrial fibrillation: a global registry program on long-term oral antithrombotic treatment in patients with atrial fibrillation. *Am Heart J* 2014;**167**:329–334.
20. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, *et al.* 2016 ESC guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016;**37**:2893–2962.
21. Wynn GJ, Todd DM, Webber M, Bonnett L, McShane J, Kirchhof P, *et al.* The European Heart Rhythm Association symptom classification for atrial fibrillation: validation and improvement through a simple modification. *Eur Eur Pacing Arrhythm Card Electrophysiol J Work Groups Card Pacing Arrhythm Card Cell Electrophysiol Eur Soc Cardiol* 2014;**16**:965–972.
22. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. *J Stat Softw* 2011;**45**:1–67.
23. van der Wall SJ, Lip GYH, Teutsch C, Kalejs O, Lyrer P, Hall C, *et al.* Low bleeding and thromboembolic risk with continued dabigatran during cardiovascular interventions: the GLORIA-AF study. *Eur J Intern Med* 2021;**91**:75–80.
24. Saarela M, Jauhiainen S. Comparison of feature importance measures as explanations for classification models. *SN Appl Sci* 2021;**3**:272.
25. Breiman L. Random forests. *Mach Learn* 2001;**45**:5–32.
26. Gorog DA, Gue YX, Chao T-F, Fauchier L, Ferreiro JL, Huber K, *et al.* Assessment and mitigation of bleeding risk in atrial fibrillation and venous thromboembolism: executive summary of a European and Asia-Pacific expert consensus paper. *Thromb Haemost* 2022;**122**:1625–1652.
27. Navar AM, Kolkailah AA, Overton R, Shah NP, Rousseau JF, Flaker GC, *et al.* Trends in oral anticoagulant use among 436 864 patients with atrial fibrillation in community practice, 2011 to 2020. *J Am Heart Assoc* 2022;**11**:e026723.
28. Fox KAA, Virdone S, Pieper KS, Bassand J-P, Camm AJ, Fitzmaurice DA, *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**:214–227.
29. Lip GYH. The ABC pathway: an integrated approach to improve AF management. *Nat Rev Cardiol* 2017;**14**:627–628.
30. Romiti GF, Proietti M, Bonini N, Ding WY, Boriani G, Huisman MV, *et al.* Adherence to the atrial fibrillation better care (ABC) pathway and the risk of major outcomes in patients with atrial fibrillation: a post-hoc analysis from the prospective GLORIA-AF registry. *eClinicalMedicine* 2023;**55**:101757.
31. Romiti GF, Pastori D, Rivera-Caravaca JM, Ding WY, Gue YX, Menichelli D, *et al.* Adherence to the 'atrial fibrillation better care' pathway in patients with atrial fibrillation: impact on clinical outcomes – a systematic review and meta-analysis of 285,000 patients. *Thromb Haemost* 2022;**122**:406–414.
32. Huisman MV, Rothman KJ, Paquette M, Teutsch C, Diener H-C, Dubner SJ, *et al.* Two-year follow-up of patients treated with dabigatran for stroke prevention in atrial fibrillation: global registry on long-term antithrombotic treatment in patients with atrial fibrillation (GLORIA-AF) registry. *Am Heart J* 2018;**198**:55–63.
33. Chao T-F, Wang K-L, Liu C-J, Lin Y-J, Chang S-L, Lo L-W, *et al.* Age threshold for increased stroke risk among patients with atrial fibrillation: a nationwide cohort study from Taiwan. *J Am Coll Cardiol* 2015;**66**:1339–1347.
34. Mazurek M, Shantsila E, Lane DA, Wolff A, Proietti M, Lip GYH. Secondary versus primary stroke prevention in atrial fibrillation: insights from the Darlington atrial fibrillation registry. *Stroke* 2017;**48**:2198–2205.
35. Olier I, Ortega-Martorell S, Pieroni M, Lip GYH. How machine learning is impacting research in atrial fibrillation: implications for risk prediction and future management. *Cardiovasc Res* 2021;**117**:1700–1717.
36. Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, *et al.* Cardiovascular event prediction by machine learning. *Circ Res* 2017;**121**:1092–1101.
37. Li X, Liu H, Du X, Zhang P, Hu G, Xie G, *et al.* Integrated machine learning approaches for predicting ischemic stroke and thromboembolism in atrial fibrillation. *AMIA Annu Symp Proc* 2017;**2016**:799–807.
38. Goto S, Pieper KS, Bassand J-P, Camm AJ, Fitzmaurice DA, Goldhaber SZ, *et al.* New artificial intelligence prediction model using serial prothrombin time international normalized ratio measurements in atrial fibrillation patients on vitamin K antagonists: GARFIELD-AF. *Eur Heart J—Cardiovasc Pharmacother* 2020;**6**:301–309.
39. Chao T-F, Chan Y-H, Chiang C-E, Tuan T-C, Liao J-N, Chen T-J, *et al.* Early rhythm control and the risks of ischemic stroke, heart failure, mortality, and adverse events when performed early (<3 months): a nationwide cohort study of newly diagnosed patients with atrial fibrillation. *Thromb Haemost* 2022;**122**:1899–1910.
40. Romiti GF, Proietti M, Bonini N, Ding WY, Boriani G, Huisman MV, *et al.* Clinical complexity domains, anticoagulation, and outcomes in patients with atrial fibrillation: a report from the GLORIA-AF registry phase II and III. *Thromb Haemost* 2022;**122**: 2030–2041.