Use of Machine Learning and Statistical Inference Methods for Identification of Risk Factors Associated with Atrial Fibrillation in Indian Patients: A Real-World Retrospective Study

Namrata Kulkarni, Santosh Taur, Danai Aristeridou, Salil Shinde, Konstantinos Spyridopoulos, Ahsan Huda, Sonali Dighe

Submitted to: JMIR Cardio
on: June 26, 2024

Table of Contents

Use of Machine Learning and Statistical Inference Methods for Identification of Risk Factors Associated with Atrial Fibrillation in Indian Patients: A Real-World Retrospective Study

Namrata Kulkarni[1]; Santosh Taur[2]; Danai Aristeridou[3*]; Salil Shinde[4*]; Konstantinos Spyridopoulos[3*]; Ahsan Huda[5*]; Sonali Dighe[2*]

[1]Department of Medical Affairs, Pfizer Limited Mumbai IN
[2]Department of Medical Affairs, Pfizer Limited, Plot No. C, The Capital, 1802, 18th Floor, 70, G Block Road, Bandra Kurla Complex, Bandra East, Mumbai- 400051, Maharashtra, India Mumbai IN
[3]Department of Data Science, Pfizer Limited, Leoforos Georgikis Scholis 27, Thermi Thessaloniki 57001, Greece Thermi Thessaloniki GR
[4]Department of Medical Affairs, Pfizer Corporation, Kerry Center, 683, King's Rd, Quarry Bay, Hong Kong, China Hong Kong HK
[5]Department of Data Science, Pfizer Limited, 235 E. 42nd Street, New York, New York- 10017, USA New York US
[*]these authors contributed equally

Corresponding Author:
Namrata Kulkarni
Department of Medical Affairs, Pfizer Limited
Plot No. C, The Capital, 1802, 18th Floor, 70, G Block Rd, Bandra Kurla Complex, Bandra East, Mumbai, Maharashtra 400051, India
Mumbai
IN

Abstract

Background: Atrial Fibrillation (AF) is the most common type of cardiac arrhythmia in the general population. Machine Learning (ML) models can be utilized to identify the most pertinent risk factors and build reliable risk factor-based screening algorithms.

Objective: The aim of current study was to evaluate ML in identifying risk factors associated with AF in Indian patients in tertiary care settings.

Methods: This real-world, retrospective, observational, multicenter study was conducted by collecting data from anonymized Electronic Medical Records (EMRs) of patients between 1 Jan 2016 and 31 Dec 2019. The ML logistic regression model was used to calculate the Odds Ratio (OR) for identifying the risk factors.

Results: A total of 5044 patients were included in the study; cohort of AF (cases; n=2516) and non-AF patients (controls; n=2528). The OR showed age above 65 years and female gender as risk factors for incident AF. Smoking, alcohol consumption, dyslipidemia, and hemiplegia, were strongly associated risk factors for AF with high ORs. Chronic diseases such as cardiac arrhythmia, paroxysmal tachycardia, rhematic heart disease, valvular heart disease, chronic obstructive pulmonary disease, stroke, surgery, neurological history, and hypothyroidism also emerged from the model as risk factors for developing incident AF. The model had an F1-score of 0.6930, predicting incident AF with an accuracy of 70.27% (accuracy=0.7027), sensitivity of 0.7042, and specificity of 0.7013. The model also had a positive predictive value of 0.6820 and a negative predictive value of 0.7230 with a threshold of 0.51 for the classification of true AF patients and an area under curve (AUC) of 0.7756.

Conclusions: The ML logistic regression model could be a useful predictive tool for incident AF that aids in AF risk management and individualized clinical decision-making.

(JMIR Preprints 26/06/2024:63662)
DOI: https://doi.org/10.2196/preprints.63662

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Use of Machine Learning and Statistical Inference Methods for Identification of Risk Factors Associated with Atrial Fibrillation in Indian Patients: A Real-World Retrospective Study

Namrata Kulkarni,[1] Santosh Taur,[1] Danai Aristeridou,[2] Salil Shinde,[3] Konstantinos Spyridopoulos,[2] Ahsan Huda,[4] Sonali Dighe[1]

[1] Department of Medical Affairs, Pfizer Limited, Plot No. C, The Capital, 1802, 18th Floor, 70, G Block Road, Bandra Kurla Complex, Bandra East, Mumbai- 400051, Maharashtra, India

[2] Department of Data Science, Pfizer Limited, Leoforos Georgikis Scholis 27, Thermi Thessaloniki 57001, Greece

[3] Department of Medical Affairs, Pfizer Corporation, Kerry Center, 683, King's Rd, Quarry Bay, Hong Kong, China

[4] Department of Data Science, Pfizer Limited, 235 E. 42nd Street, New York, New York- 10017, USA

Correspondence should be addressed to Namrata Kulkarni; namrata.kulkarni@pfizer.com

Abstract

Background: Atrial Fibrillation (AF) is the most common type of cardiac arrhythmia in the general population. Machine Learning (ML) models can be utilized to identify the most pertinent risk factors and build reliable risk factor-based screening algorithms.

Objective: The aim of current study was to evaluate ML in identifying risk factors associated with AF in Indian patients in tertiary care settings.

Methods: This real-world, retrospective, observational, multicenter study was conducted by collecting data from anonymized Electronic Medical Records (EMRs) of patients between 1 Jan 2016 and 31 Dec 2019. The ML logistic regression model was used to calculate the Odds Ratio (OR) for identifying the risk factors.

Result: A total of 5044 patients were included in the study; cohort of AF (cases; n=2516) and non-AF patients (controls; n=2528). The OR showed age above 65 years and female gender as risk factors for incident AF. Smoking, alcohol consumption, dyslipidemia, and hemiplegia, were strongly associated risk factors for AF with high ORs. Chronic diseases such as cardiac arrhythmia, paroxysmal tachycardia, rhematic heart disease, valvular heart disease, chronic obstructive pulmonary disease, stroke, surgery, neurological history, and hypothyroidism also emerged from the model as risk factors for developing incident AF. The model had an F1-score of 0.6930, predicting incident AF with an accuracy of 70.27% (accuracy=0.7027), sensitivity of 0.7042, and specificity of 0.7013. The model also had a positive predictive value of 0.6820 and a negative predictive value of 0.7230 with a threshold of 0.51 for the classification of true AF patients and an area under curve (AUC) of 0.7756.

Conclusion: The ML logistic regression model could be a useful predictive tool for incident AF that aids in AF risk management and individualized clinical decision-making.

Keywords: Atrial fibrillation; machine learning; risk factors; logistic regression model

Introduction

Atrial Fibrillation (AF) is a condition characterized by supraventricular tachyarrhythmia with uncoordinated atrial activation and subsequently futile atrial contraction [1]. AF is the most common prolonged cardiac arrhythmia in the general population [2]. In the last 50 years, the prevalence of AF has increased 3-fold based on data from the Framingham Heart Study (FHS) [3]. As per the Global Burden of Disease (GBD), 37.57 million prevalent cases and 3.05 million incident cases of AF were reported worldwide, contributing to 287,241 deaths [4].

Atrial Fibrillation is a multifaceted condition varying from an isolated electrophysiological disorder to a manifestation or outcome of other cardiac and noncardiac pathologies [5,6]. Stroke is the most concerning AF complication and AF is associated with a 4-to-5-fold increased risk of stroke [3]. The relevance of AF in India might be evaluated considering the likelihood that 20% of strokes among India's 0.4 million stroke victims might be related to AF [7]. A study from Ludhiana evaluated 1,942 patients with stroke and found AF in 203 (10%) patients [8]. Another hospital-based study in Dehradun, reported 62/246 (25%) patients with stroke and AF [9]. However, there are no large population-based prevalence studies of AF conducted in India.

The targeted AF screening approach has been successfully implemented in high-income countries [10-12] However, low- and middle-income countries, including India, lack such systematic efforts to screen for AF. With Machine Learning (ML) techniques, we can select the most pertinent risk factors from high-dimensional data and build reliable risk factor-based screening algorithms [13], which can help identify non-linear associations and complex interactions between variables without requiring pre-specified relationships. A recent study reported that ML algorithms are accurate in detecting AF by transmitting electrocardiogram (ECG) data, enabling confirmation of AF, and improving the identification of newly diagnosed AF patients at risk [14]. The Cohorts for Heart and Aging Research in Genomic Epidemiology for Atrial Fibrillation (CHARGE-AF) have developed a ML model which delivers a simple 5-year AF risk [15]. Thus, through ML, we can predict diseases that would enable physicians to make quick decisions, by expanding their efficiency, and improving their diagnostic accuracy.

However, most existing models use variables such as ECG parameters that are not commonly available in the Electronic Medical Records (EMR). There have been attempts to develop EMR-based models, but these required either the use of lengthy prediction horizons or data harmonization as an additional step [16]. Other models that use data from the EMRs produce a 5-year risk prediction using binned risk categories [17, 18]. The ML AF-risk model may lower the risk of stroke or systemic embolism and aid healthcare practitioners in understanding the characteristics of the disease. Thus, the present study aimed to develop a ML model from EMR data to identify risk factors associated with the development of incident AF in Indian patients in tertiary care settings.

Methods

Study design and data sources

This real-world, retrospective, observational, multicenter study was conducted by collecting patients' data from anonymized EMRs between 1 Jan 2016 and 31 Dec 2019 from seven tertiary care centers including private hospitals, institutes, and medical colleges across India. The baseline period was from 01 Jan 2016 to 31 Dec 2017 for patient identification followed by an outcome period (01 Jan 2018 to 31 Dec 2019) and an index date—the first day of AF recording. The ML logistic regression model was used to calculate the odds ratio (OR) for identifying the risk factors.

Study population

The study inclusion criteria were patients aged ≥45 years, with a continuous enrolment in the dataset for 24 months pre-index date (baseline period: 01 Jan 2016 – 31 Dec 2017) and with incident documented AF diagnosis between 01 Jan 2018 and 31 Dec 2019 (outcome period) for cases and till the end of study (01 Jan 2016 to 31 Dec 2019) or death for controls. Patients had to have two ECG-datapoints (preferably one each year) during the baseline period, and medical history present in the EMR with diagnosis, clinical visits, medical prescriptions, and administration or any other recorded activity. For cases, it was required to have a diagnosis of AF and a new ambulatory verified or primary or secondary hospital discharge diagnosis of AF during the outcome period. Patients with AF diagnosis in the 24 months pre-index period, with unknown status of AF during the outcome period, and with incomplete EMRs were excluded.

Ethical considerations

The study was conducted as per the protocol and principles of the Declaration of Helsinki and in accordance with legal and regulatory requirements, as well as with scientific purpose, value, and rigor, and followed generally accepted research practices described in International Ethical Guidelines for Epidemiological Studies issued by the Council for International

Organizations of Medical Sciences (CIOMS) [19]. The study was conducted after approval by the Royal Pune Independent Ethics Committee, Pune; Institutional Ethics Committee (IEC), Fortis Hospital, Mohali; IEC, Bio Medical Research Apollo Hospitals, Chennai; IEC, PD Hinduja Hospital and Medical Research Centre, Mumbai; IEC, Sehgal Nursing Home, Delhi; and IEC, Gleneagles Global Hospitals, Hyderabad. As it was a retrospective study involving EMRs of patients, the need for informed consent was waived by the ethics committee.

Statistical analysis

Data analysis was performed internally. Data preparation for the analysis was done by standardizing data records, re-coding variables, creating the derived categorical variables from continuous variables, and removing the missing values for the variables (If a variable had high proportion of missing values i.e., threshold set to 20%, the variables were not used for the analysis). For the feature selection or selection of the significant variables for the logistic model, variables with less than 5% prevalence in each class, considered as having low support, were dropped. The Variance Inflation Factor (VIF) was used to assess the multicollinearity and Partial Least Squares (PLS) regression was used to assess the potential to support feature selection. The exploratory data analysis for cases (n=2516) and controls (n=2528) was carried out. The categorical variables were presented by n (%) and the numerical variables were presented by mean ± standard deviation (SD). A stepwise logistic regression model was used on the set of variables that were all statistically significant from the feature selection approaches. This is also known as a parsimonious approach. The dataset was further split into two data subsets (80% for training and 20% for testing). On every iteration, the model dropped a variable with p-value higher than 0.05 and repeated until all the remaining variables were statistically significant ($p<0.05$). All the logistic regression models, even the preliminary, were fine-tuned by experimenting with different optimizers (e.g., newton, bfgs, powell, cg, ncg, basinhopping, minimize). The assessment of the performance of the regression model was done based on the optimal cutoff and the diagnostic accuracy parameters [overall accuracy, f1 score, sensitivity, specificity, positive and negative predictive and, area under curve and Receiver operating characteristic (AUC-ROC)]. Ethical AI analysis was used to obtain the OR for the selected variables to assess the effect of the selected explanatory variables on the outcome variable (development of AF). The explanatory variables with OR above 1.0 were considered as risk factors and below 1.0 were considered as protective factors.

Results

Demographic and baseline characteristics

A total of 5044 patient records were screened; most patients were in the age group of 55-64 years in both control [970 (38.37%)] and case [898 (35.69%)] groups with majority of patients being males in both control [1751 (69.26%)] and case groups [1562 (62.08%)]. Further, hypertension was reported by 1242 (49.13%) patients in the control group and 948 (37.68%) patients in the case group. Obesity was reported in 259 (10.29%) patients in the control group and 133 (5.26%) patients in the case group (see Table 1). The other baseline characteristics Body Mass Index (BMI), pulse and blood pressure, lipid parameters, thyroid parameters, and glycemic parameters are presented in the Supplementary Tables S-1 and S-2.

Table *1*: Demographic and baseline characteristics of patients (N=5044)

Variables	**Control (N=2528)**	**Case (N=2516)**
	n (%)	**n (%)**
Age Distribution; n (%)		
45-54 Years	290 (11.47%)	211 (8.39%)
55-64 Years	970 (38.37%)	898 (35.69%)
65-74 Years	707 (27.97%)	777 (30.88%)
75-84 Years	456 (18.04%)	513 (20.39%)
≥85 Years	105 (4.15%)	117 (4.65%)
Gender, n (%)		
Female	777 (30.74%)	954 (37.92%)
Male	1751 (69.26%)	1562 (62.08%)
Hypertension (mmHg), n (%)	1242 (49.13%)	948 (37.68%)
Duration since diagnosis (years) Mean ± SD	8.27 ± 5.02	4.66 ± 2.93
Obesity, n (%)	259 (10.29%)	133 (5.26%)
Duration since diagnosis (years) Mean ± SD	9.28 ± 3.27	5.03 ± 3.79

SD: Standard deviation

Telmisartan (18.43%), atorvastatin (17.96%), sulfonylureas (14%), and amlodipine (12.74%) were the most prescribed concomitant drugs in the controls, whereas aspirin (10.89%), sulfonylureas (10.02%), telmisartan (9.62%), and amlodipine/atorvastatin (7.47%/7.35%) were the most prescribed concomitant drugs in the cases (Supplementary Table S-3). Baseline comorbidities are presented in Supplementary Table S-4.

Risk of incident AF

The below-mentioned findings demonstrated gender male, age 55 - ≥85 years, chronic obstructive pulmonary disease, diabetes mellitus, Triglyceride (TG) 150-199 mg/dL, Low Density Lipoprotein (LDL) 100-129 mg/dL, High Density Lipoprotein (HDL) <40 mg/dL, HDL ≥60 mg/dL, heart failure, cardiovascular history, past surgical history, history of smoking and alcohol, hypertension, obesity, elevated blood pressure, isolated systolic hypertension have statistically significant (p-value<0.001) impact (either risk or protective) on the development of AF (see Table 2).

Table *2*: Association of independent variables on incident atrial fibrillation

Parameter	Regression Coefficient	Standard Error	z value	Regression Coefficient (95% CI)	p-value
Sex, male	-0.397	0.080	-4.982	-0.554, -0.241	<0.001
Age, 55-64 Years	0.341	0.098	3.490	0.15, 0.532	<0.001
Age, 65-74 Years	0.560	0.106	5.288	0.352, 0.767	<0.001
Age, 75-84 Years	0.540	0.113	4.774	0.318, 0.762	<0.001
Age, ≥85 Years	0.502	0.176	2.856	0.158, 0.847	0.004
Chronic obstructive pulmonary disease	1.712	0.294	5.827	1.136, 2.287	<0.001
Diabetes mellitus	-1.338	0.087	-15.369	-1.508, -1.167	<0.001
TG 150-199 mg/dL	0.223	0.105	2.133	0.018, 0.428	0.033
LDL 100-129 mg/dL	0.625	0.119	5.257	0.392, 0.858	<0.001
HDL <40 mg/dL	1.224	0.092	13.343	1.044, 1.403	<0.001
HDL ≥60 mg/dL	0.937	0.091	10.247	0.758, 1.117	<0.001
Heart failure	-0.309	0.129	-2.392	-0.563, -0.056	0.017
Cardiovascular history	-1.725	0.105	-16.425	-1.931, -1.519	<0.001
Past surgical history	1.273	0.196	6.510	0.89, 1.657	<0.001
History of smoking	1.056	0.263	4.021	0.541, 1.571	<0.001
History of alcohol	1.997	0.312	6.405	1.386, 2.608	<0.001
Hypertension	-0.194	0.078	-2.504	-0.346, -0.042	0.012
Obesity	-0.823	0.143	-5.769	-1.103, -0.543	<0.001
Elevated blood pressure	-0.966	0.129	-7.477	-1.219, -0.713	<0.001
Isolated systolic hypertension	-0.646	0.101	-6.401	-0.844, -0.448	<0.001

CI: Confidence interval, HDL: High density lipoprotein, LDL: Low density lipoprotein, TG: Triglycerides

Pseudo R-Square = 0.244; Log-Likelihood Value = -2115.30; p-value <0.001

The OR showed age above 65 years and female gender as risk factors for incident AF. Smoking, alcohol consumption, dyslipidemia, and hemiplegia, were strongly associated risk factors for AF with high ORs. Chronic diseases such as cardiac arrhythmia, paroxysmal tachycardia, rhematic heart disease, valvular heart disease, chronic obstructive pulmonary disease, stroke, surgery, orthopedic history, neurological history, and hypothyroidism also emerged from the model as risk factors for developing incident AF (see Table 3 and Figure 1).

The OR was less than 1 for adverse drug reaction, age 45-54 years, Diabetes Mellitus (DM), obesity, elevated blood pressure, myocardial infarction, hypertension, coronary heart disease and renal disease suggesting that these variables are not the risk factors for developing incident AF (see Table 3 and Figure 1).

Table *3*: Impact of independent variables on incident atrial fibrillation

Variables	**Odds Ratio (95% CI)**
Adverse drug reaction	0 (0, 0)
Age 55-64*	1.0
Age 45-54	0.7 (0.6, 1)
Age 65-74	1.1 (1, 1.4)
Age 75-84	1.2 (1.1, 1.4)
Age ≥85	1.1 (0.9, 1.5)
Female gender	1.3 (1.2, 1.5)
High blood pressure	1.0
Elevated blood pressure	0.4 (0.4, 0.6)
Normal blood pressure	1.0 (0.9, 1.2)
Cardiac arrhythmia	1.6 (0.9, 2.9)
Cardiovascular history	0.2 (0.2, 0.2)
COPD	6.8 (4.5, 10.4)
Coronary heart disease	0.4 (0.3, 0.7)
Diabetes mellitus	0.2 (0.3, 0.3)
Hemiplegia	10.0 (5.5, 21.6)
HDL_≥ 80 mg/dL (High*)	1.0
HDL ≥60 mg/dL (Good)	2.2 (1.9, 2.5)
HDL <40 mg/dL (Low)	3.1 (2.8, 3.7)
LDL <100 mg/dL	1.0
LDL 100-129 mg/dL	1.4 (1.2, 1.7)
LDL 130-159 mg/dL	1.6 (1.2, 2.4)
LDL 160-189 mg/dL	2.9 (0.9, 9.2)
LDL ≥190 mg/dL	12.0 (1.7, 98.5)
TC <200 mg/dL*	1.0
TC 200-239 mg/dL	1.8 (1.4, 2.5)
TC ≥240 mg/dL	7.8 (3.9, 15.7)
TG <150 mg/dL*	1.0
TG 150-199 mg/dL	0.9 (0.8, 1.1)
TG 200-499 mg/dL	9.9 (4.6, 21.7)
History of alcohol	11.0 (7.0, 18.4)
History of smoking	5.4 (3.7, 8.2)
Hypertension	0.6 (0.6, 0.7)
Myocardial infraction	0.9 (0.8, 1.1)
Neurological history	2.3 (0.9, 6.1)
Obesity	0.4 (0.4, 0.6)
Obstructive respiratory disorder	0.5 (0.2, 1.3)

Orthopedic history	1.6 (1.2, 2.3)
Paroxysmal tachycardia	5.8 (3.7, 9.1)
Pulse Rate 60-100 bpm*	1.0
Pulse Rate <60 bpm	0.2 (0, 1.7)
Pulse Rate >100 bpm	0.8 (0.3, 2.4)
Renal disease history	0.1 (0.1, 0.3)
Rhematic heart disease	1.4 (0.7, 3)
Hyperthyroidism	1.0 (0.1, 7.2)
Hypothyroidism	1.1 (0.9, 1.5)
Stroke	1.3 (1.1, 1.8)
Valvular heart disease	4.2 (2.6, 6.8)
Past surgical history	2.6 (2, 3.4)
CABG	Null
Cholecystectomy	4.6 (2.8, 7.7)
Hemorrhoidectomy	5.4 (3.2, 9.4)
Chronic kidney disease	1.0 (0.9, 1.3)
Family member with medical history	1.0 (0.1, 16.1)
Heart failure	1.0 (0.9, 1.3)
Chronic infectious diseases history	Null
Gynecological history	Null
Other pulmonary heart disease	Null
Respiratory history	Null

CABG: Coronary artery bypass graft, CI: Confidence interval, COPD: Chronic obstructive pulmonary disease, HDL: High density lipoprotein, LDL: Low density lipoprotein, TC: Total cholesterol, TG: Triglycerides,

[*] Note: Absence of disease was considered as reference (OR=1)

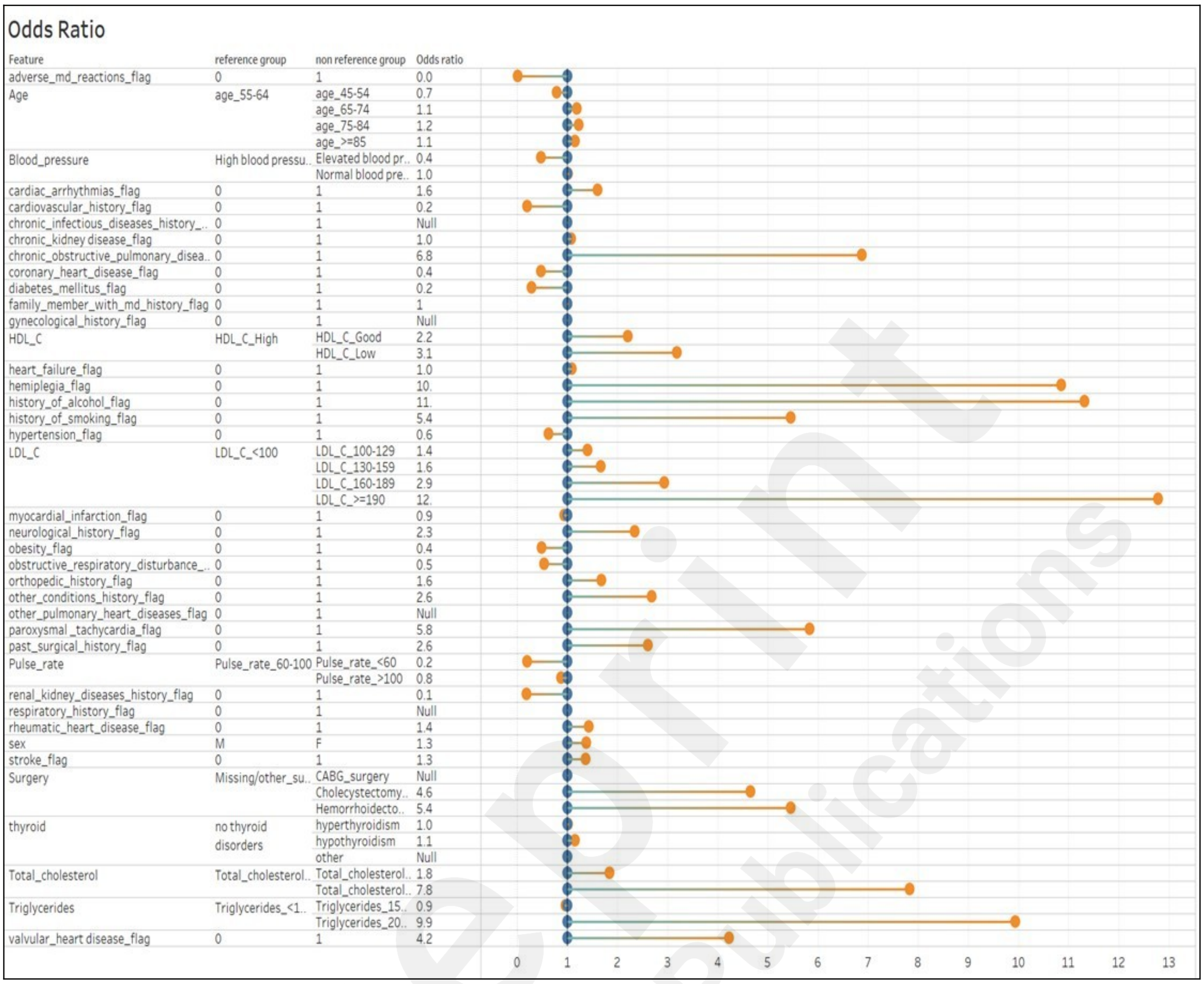

Figure 1: Forest plot of the impact of independent variables on AF

The logistic regression model reported Pseudo R-Square = 0.244 i.e., the model was able to explain 24.4% variation in the outcome variable and the Log-Likelihood Value at -2115.30. This regression model was able to fit the observed data significantly ($p<0.001$). The model had an F1-score of 0.6930, predicting incident AF with an accuracy of 70.27% (accuracy=0.7027), sensitivity of 0.7042, and specificity of 0.7013. The model also had a positive predictive value of 0.6820 and a negative predictive value of 0.7230 with a threshold of 0.51 for the classification of true AF patients and an AUC of 0.7756 (see Figure 2).

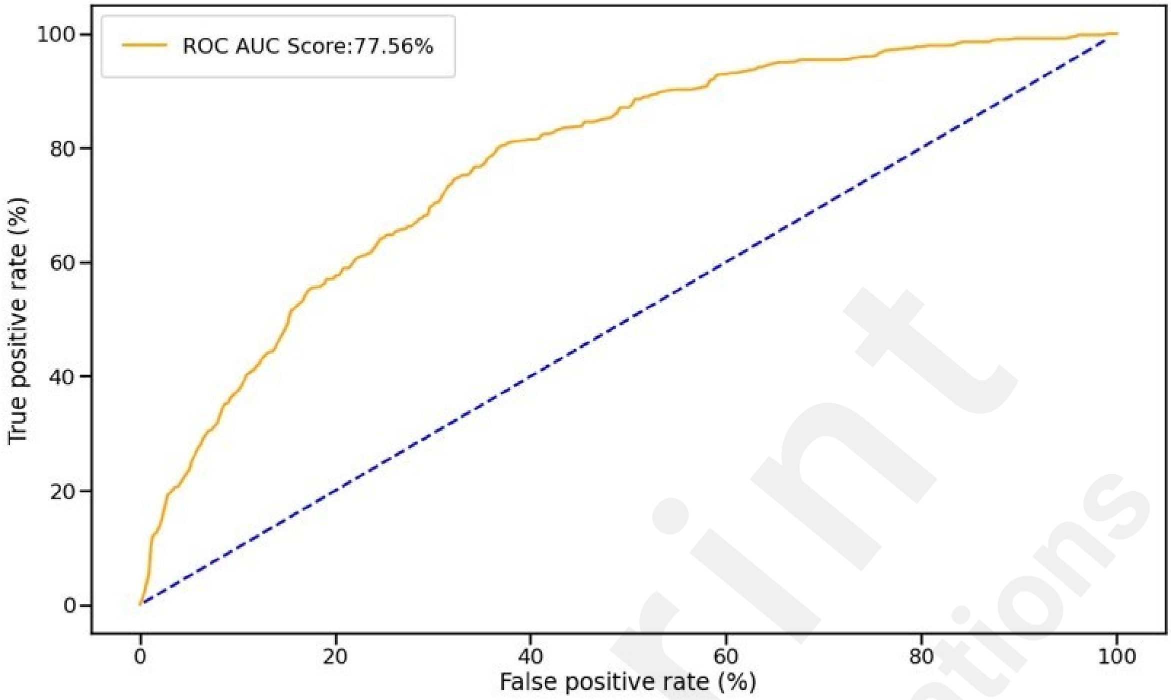

Figure 2: ROC curve

Discussion

Principle Findings

Clinical risk prediction models have been extensively used in several medical domains. By utilizing individual-level data, these models aim to predict clinically relevant outcomes. The current study identified the predictors for incident AF by using ML techniques. The OR calculated based on the ML reported that age above 65 years, female gender, alcohol consumption and smoking, hemiplegia, dyslipidemia, hypothyroidism, paroxysmal tachycardia, rheumatic heart disease, stroke, valvular heart disease, cardiac arrhythmia were the risk factors for developing incident AF.

The ML regression model used had an F1-score of 0.6930, accuracy of 70.27% (accuracy=0.7027) sensitivity of 0.7042, specificity of 0.7013, positive predictive value of 0.6820, and negative predictive value of 0.7230 with threshold of 0.51 for the classification of true AF patients and an AUC of 0.7756. The AUC value 0.7756 (AUC=77.56%) indicated that the model could distinguish the true AF patients from the population. The current regression model reported 0.244 Pseudo R-Square which measured the independent variables to explain the variation in the dependent variables. Therefore, 24.4% of the variation in the dependent variables were explained by the independent variables. Further, the current model reported a Log-Likelihood value as -2115.30. Hence, this regression model was able to fit the observed data significantly (p<0.001).

Comparison With Prior Work

Similar to the current study findings, Humel et al. had estimated AF risk using routinely ascertained features in electronic health record (EHR) and validated a prediction model for five-year AF risk with female gender, age above 65 years, smoking, hyperlipidemia, valvular heart disease, and hypothyroidism as AF risk factors [17].

The atrial myocardium undergoes electrical and structural remodeling with age, both of which may play an important role in the initiation and/or perpetuation of atrial arrhythmias and provide a substrate for AF [20]. The current study reported age above 65 years and female gender as the risk factors for developing incident AF, comparable to the results from a retrospective cohort study by Piccini et al., which reported incident AF (n=433, 123) in 55% of females and patients with a mean age of 80 years [21].

Smoking and alcohol consumption increase the lifetime risk of a rapid and irregular heart rate, and the development of AF [22]. In the current study, patients might have an unhealthy lifestyle with excess consumption of alcohol and tobacco together, which triggers a risk factor for the development of AF. Similar to current study findings, several clinical studies have suggested that smoking and high alcohol consumption are risk factors for incident AF [23-25]. Lu et al. reported that the OR per one-unit increase of smoking initiation was 1.11 (95% confidence interval (CI), 1.06–1.16; $P=3.35\times10^{-6}$) for AF, and intake of heavy alcohol increased the risk of AF (OR 1.11; 95% CI, 1.04–1.18; P=0.001) [26].

It is a known fact that dyslipidemia is a major risk factor for cardiovascular disease. Elevated LDL and decreased levels of HDL are associated with reduced systolic and diastolic left ventricle function, which are risk factors for AF [27]. This was seen in the current study's outcomes as well. Additionally, a population based-study conducted on > 65,000 adults, observed lower levels of HDL-C, high TG levels, and a high TG/HDL-C ratio to be consistently associated with a higher risk of AF over 3 decades of follow-up [28].

In accordance with the current study, Schnabel et al. reported that the patients with hemiplegia and paroxysmal tachycardia were 3.04 (95% CI 2.86–3.23) and 2.20 (95% CI 2.11–2.30) times more likely to be diagnosed with AF, respectively [29].

Hypothyroidism is known to be associated with cardiovascular risk factors, subclinical cardiovascular disease, and overt cardiovascular disease, all of which predispose to AF [30]. Similar findings were seen in the current study, where hypothyroidism was found to be a risk factor for AF. Further, by using a logistic regression model, Jamies et al. reported hypothyroidism to be significantly associated with AF [Risk Ratio 1.9 (1.05-3.8, p=0.04)] [31].

In the current study, COPD was reported as a risk factor for AF, which was also observed in Liao et al., and Konecny et al. studies, where the incidence of AF was higher in COPD patients and was significantly ($p<0.05$) associated with AF [32, 33]. In a disease such as COPD, serum C-reactive protein (CRP) levels are significantly higher in AF patients. CRP reflects an inflammatory state in coronary endothelial cells, which results in thromboembolic risk [34]. The association between AF and rhematic heart disease is well established. The current study showed the presence of rhematic heart disease to be a strong predictor for development of AF. Similar associations were seen in an observational cross-sectional study by Dhungana et al., the prevalence of AF was 36.3% in 330 cases of rhematic heart disease patients [35]. Further, many studies have also revealed the prevalence of AF in patients with rhematic heart disease, ranging from 13.9% to 43% [36-38]. AF is one of the ten potentially modifiable risk factors associated with acute stroke [39]. In the current study, stroke was reported as a risk factor for AF. Similarly, Oladiran et al., and Yaghi et al., found that AF is associated with the risk of stroke, with its incidence being high in patients with AF [40, 41]. Further, in line with the current study findings, several studies have reported valvular heart disease, surgery, neurological and orthopedic history as risk factors for AF [40, 42-44]. Ambale et al. had used ML for the detection of AF using wearable technology and reported that ML improved the prediction accuracy of cardiovascular event prediction in an initially asymptomatic population [45]. Tiwari et al. had also studied ML approaches to EHR data for the identification of risk factors for AF and discovered it to be a promising method for improving risk prediction for incident AF [16].

Contrary to the current study findings, few studies have reported male sex, weight, diabetes mellitus, hypertension, heart failure, and cardiovascular history as risk factors for AF [16, 17, 46].

There is conflicting data as to whether gender plays a role in the association between various risk factors and the development of AF. Females with AF tend to have a higher incidence of valvular heart disease, while males have a higher incidence of coronary artery disease [47]. Population studies have shown a variable range of gender differences, including a higher prevalence of AF in males and a higher risk of AF recurrence in females [48]. In the present study, most patients were male in both control (69.26%) and case (62.08%) groups, which may be a confounding factor for the female gender to emerge as a risk factor for developing AF in the current study.

An association between obesity and AF has already been established in the literature. However, a 12-month longitudinal observational study by Rodriguez-Reyes et al., on patients with documented AF demonstrated that mortality in these patients was inversely associated with high BMI [49]. Post hoc analysis studies have also found that the presence of both overweight and obesity in patients with non-valvular AF is associated with a lower risk for stroke and all-cause death [50, 51]. Obese individual patients with coronary heart disease or pre-existing heart failure have a more favorable prognosis; known as the obesity paradox [52]. In addition, the current study reported obesity only in 10.29% of patients in controls and 5.26% of patients in cases, which could be another reason for obesity emerging as a protective effect for developing AF.

Previous evidence on the association of diabetes with AF symptoms has been sparse. The original Framingham heart study results showed hypertension, diabetes, congestive heart failure, and valvular heart disease as an independent risk factors for AF [47]. A survey in China had found that DM was not an independent risk factor for AF in the multivariate model [53]. Patients with DM perceive AF symptoms less often than those without DM [54, 55], with few studies suggesting no association of DM with AF and demonstrating DM as a non-significant risk predictor for AF [56-58]. Several mechanisms can explain the reduced manifestation of AF symptoms in patients with DM, a potential mechanism through which DM can affect AF symptoms is cardiac autonomic neuropathy. Diabetes-induced cardiac autonomic neuropathy can reduce the sensitivity of cardiac nerves and ultimately attenuate the perception of AF symptoms [59, 60]. Hence, the currently developed model does not indicate DM as a risk factor for developing incident AF.

AF is associated with a range of cardiac outcomes. The present model did not indicate cardiovascular history as a risk factor for developing incident AF. A large cohort study suggested that AF is not strongly associated with ischemic heart disease [61]. Instead, AF is associated with impaired coronary flow and diminished myocardial perfusion [62]. Several studies have also found the associations of AF with cardiovascular outcomes [63, 64]. In contrast, He et al., observed that the yearly risk of myocardial infarction (MI) in patients with AF was low, with pharmacological intervention skewed towards the degree of its impact [65]. Since in the current study, angiotensin-II receptor blockers, calcium channel blockers, and statins were commonly prescribed drugs and could have been a protective factor against adverse cardiovascular outcomes, the currently developed model did not indicate cardiovascular history as a risk factor for developing incident AF.

A correlation between hypertension and AF has already been demonstrated in several studies

[66, 67]. However, in line with current study findings, Thacker et al., in a multicentric population-based inception cohort study, had reported no association of hypertension with permanent AF [68]. Similar findings were reported in the Canadian Registry of Atrial Fibrillation (CARAF) study, where hypertension was not associated with AF [69]. In the present study, most of the patients with hypertension were in control (49.13%) group which may be a confounding factor for the hypertension to emerge as a protective effect for developing AF. Further, an angiotensin II receptor antagonist (telmisartan) was the most prescribed concomitant drug which could show a protective effect against AF.

Limitations

The study limitations include the inability to identify participants to predict the risk of AF with incomplete or no previous clinical data records. There was a possibility of undiagnosed AF patients ending up in the control group as they were yet to receive the right diagnosis, which would obscure the results. Thus, further studies on proof of concept are needed to implement these models in existing databases of tertiary care centers.

Conclusions

The logistic regression model used in the current study successfully identified age above 65 years, female gender, alcohol consumption and smoking, hemiplegia, dyslipidemia, hypothyroidism, paroxysmal tachycardia, rhematic heart disease, stroke, valvular heart disease, cardiac arrhythmia, COPD, surgery, neurological, and orthopedic history as the risk factors associated with the development of incident AF in Indian patients in the tertiary care settings. However, this model predicted a few risks factors contrary to other studies, which may have been due to demographics or usage of chronic disease medications by patients. The current ML model has the potential to be a valuable predictive tool for incident AF, assisting clinicians in AF risk management and individualized clinical decision-making.

Acknowledgments

The authors would like to thank Dr. Farah Iram, Rupali Jangid and Dr. Venugopal Madhusudhana from THB c/o Sekhmet Technologies Pvt. Ltd. for their contribution to manuscript writing and Mr. Raman Gupta for statistical analysis. The study was funded by Pfizer Limited.

Conflicts of Interest

All the authors are full-time employees of Pfizer Limited.

Data Availability

The statistical and image data used to support the findings of this study are included within the article and supplementary material file.

Supplementary Materials

The baseline characteristics of patients, baseline laboratory parameters, concomitant medication information and baseline comorbidities are presented in Table S1, Table S2, Table S3 and Table S4, respectively of the supplementary material file.

Abbreviations

AF: Atrial Fibrillation

AUC: Area Under Curve

BMI: Body Mass Index

CABG: Coronary Artery Bypass Graft

CARAF: Canadian Registry of Atrial Fibrillation

CHARGE-AF: Cohorts for Heart and Aging Research in Genomic Epidemiology for Atrial Fibrillation

CI: Confidence Interval

CIOMS: Council for International Organizations of Medical Sciences

CRP: C-Reactive Protein

DM: Diabetes Mellitus

ECG: Electrocardiogram

EMR: Electronic Medical Record

FHS: Framingham Heart Study

GBD: Global Burden of Disease

HDL: High Density Lipoprotein

HER: Electronic Health Record

IEC: Institutional Ethics Committee

LDL: Low Density Lipoprotein

MI: Myocardial Infarction

ML: Machine Learning

OR: Odds Ratio

PLS: Partial Least Squares

ROC: Receiver operating characteristic

SD: Standard Deviation

TC: Total Cholesterol

TG: Triglyceride

VIF: Variance Inflation Factor

References

1. Lip GY, Tse HF. Management of atrial fibrillation. Lancet. 2007 Aug 18;370(9587):604-18. doi: 10.1016/S0140-6736(07)61300-2.
2. Modi S, Modi R. Atrial fibrillation in India: is it a tide rising or a tsunami? Austin J Cardiovasc Dis Atheroscler. 2017;4(1):1030-1.
3. Kornej J, Börschel CS, Benjamin EJ, Schnabel RB. Epidemiology of Atrial Fibrillation in the 21st Century: Novel Methods and New Insights. Circ Res. 2020 Jun 19;127(1):4-20. doi: 10.1161/CIRCRESAHA.120.316340.
4. Dai H, Zhang Q, Much AA, Maor E, Segev A, Beinart R, Adawi S, Lu Y, Bragazzi NL, Wu J. Global, regional, and national prevalence, incidence, mortality, and risk factors for atrial fibrillation, 1990-2017: results from the Global Burden of Disease Study 2017. Eur Heart J Qual Care Clin Outcomes. 2021 Oct 28;7(6):574-582. doi: 10.1093/ehjqcco/qcaa061.
5. Andrade J, Khairy P, Dobrev D, Nattel S. The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms. Circ Res. 2014 Apr 25;114(9):1453-68. doi: 10.1161/CIRCRESAHA.114.303211.
6. Saggu DK, Sundar G, Nair SG, Bhargava VC, Lalukota K, Chennapragada S, Narasimhan C, Chugh SS. Prevalence of atrial fibrillation in an urban population in India: the Nagpur pilot study. Heart Asia. 2016 Apr 18;8(1):56-9. doi: 10.1136/heartasia-2015-010674.
7. Pandian JD, Sudhan P. Stroke epidemiology and stroke care services in India. J Stroke. 2013 Sep;15(3):128-34. doi: 10.5853/jos.2013.15.3.128.
8. Akanksha WG, Paramdeep K, Gagandeep S, Rajinder B, Birinder SP, Monika S, Shavinder S, Clarence JS, Shweta JV, Sharma M, Jeyaraj DP. Clinical Features, Risk Factors, and Short-term Outcome of Ischemic Stroke, in Patients with Atrial Fibrillation: Data from a Population-based Study. Ann Indian Acad Neurol. 2017 Jul-Sep;20(3):289-293. doi: 10.4103/aian.AIAN_16_17.
9. Goel D, Gupta R, Keshri T, Rana S. Prevalence of atrial fibrillation in acute ischemic stroke patients: A hospital-based study from India. Brain Circ. 2020 Feb 18;6(1):19-

25. doi: 10.4103/bc.bc_19_19.

10. Lowres N, Krass I, Neubeck L, Redfern J, McLachlan AJ, Bennett AA, Freedman SB. Atrial fibrillation screening in pharmacies using an iPhone ECG: a qualitative review of implementation. Int J Clin Pharm. 2015 Dec;37(6):1111-20. doi: 10.1007/s11096-015-0169-1.
11. Gwynne K, Flaskas Y, O'Brien C, Jeffries TL, McCowen D, Finlayson H, Martin T, Neubeck L, Freedman B. Opportunistic screening to detect atrial fibrillation in Aboriginal adults in Australia. BMJ Open. 2016 Nov 15;6(11):e013576. doi: 10.1136/bmjopen-2016-013576.
12. Moran PS, Flattery MJ, Teljeur C, Ryan M, Smith SM. Effectiveness of systematic screening for the detection of atrial fibrillation. Cochrane Database Syst Rev. 2013 Apr 30;(4):CD009586. doi: 10.1002/14651858.CD009586.pub2. Update in: Cochrane Database Syst Rev. 2016 Jun 03;(6):CD009586. doi: 10.1002/14651858.CD009586.pub3.
13. Schnabel RB, Witt H, Walker J, Ludwig M, Geelhoed B, Kossack N, Schild M, Miller R, Kirchhof P. Machine learning-based identification of risk-factor signatures for undiagnosed atrial fibrillation in primary prevention and post-stroke in clinical practice. Eur Heart J Qual Care Clin Outcomes. 2022 Dec 13;9(1):16-23. doi: 10.1093/ehjqcco/qcac013.
14. Lown M, Brown M, Brown C, Yue AM, Shah BN, Corbett SJ, Lewith G, Stuart B, Moore M, Little P. Machine learning detection of Atrial Fibrillation using wearable technology. PLoS One. 2020 Jan 24;15(1):e0227401. doi: 10.1371/journal.pone.0227401.
15. Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB, Sinner MF, Sotoodehnia N, Fontes JD, Janssens AC, Kronmal RA, Magnani JW, Witteman JC, Chamberlain AM, Lubitz SA, Schnabel RB, Agarwal SK, McManus DD, Ellinor PT, Larson MG, Burke GL, Launer LJ, Hofman A, Levy D, Gottdiener JS, Kääb S, Couper D, Harris TB, Soliman EZ, Stricker BH, Gudnason V, Heckbert SR, Benjamin EJ. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. J Am Heart Assoc. 2013 Mar 18;2(2):e000102. doi: 10.1161/JAHA.112.000102.
16. Tiwari P, Colborn KL, Smith DE, Xing F, Ghosh D, Rosenberg MA. Assessment of a Machine Learning Model Applied to Harmonized Electronic Health Record Data for the Prediction of Incident Atrial Fibrillation. JAMA Netw Open. 2020 Jan

3;3(1):e1919396. doi: 10.1001/jamanetworkopen.2019.19396.

17. Hulme OL, Khurshid S, Weng LC, Anderson CD, Wang EY, Ashburner JM, Ko D, McManus DD, Benjamin EJ, Ellinor PT, Trinquart L, Lubitz SA. Development and Validation of a Prediction Model for Atrial Fibrillation Using Electronic Health Records. JACC Clin Electrophysiol. 2019 Nov;5(11):1331-1341. doi: 10.1016/j.jacep.2019.07.016.
18. Daly B, Gorenshteyn D, Nicholas KJ, Zervoudakis A, Sokolowski S, Perry CE, Gazit L, Baldwin Medsker A, Salvaggio R, Adams L, Xiao H, Chiu YO, Katzen LL, Rozenshteyn M, Reidy-Lagunes DL, Simon BA, Perchick W, Wagner I. Building a Clinically Relevant Risk Model: Predicting Risk of a Potentially Preventable Acute Care Visit for Patients Starting Antineoplastic Treatment. JCO Clin Cancer Inform. 2020 Mar;4:275-289. doi: 10.1200/CCI.19.00104.
19. International Ethical Guidelines for Epidemiological Studies. [Internet]. Available from: Available at https://cioms.ch/shop/product/international-ethical-guidelines-for-epidemiologicalstudies/ Accessed on 25/05/2021.
20. Wasmer K, Eckardt L, Breithardt G. Predisposing factors for atrial fibrillation in the elderly. J Geriatr Cardiol. 2017 Mar;14(3):179-184. doi: 10.11909/j.issn.1671-5411.2017.03.010.
21. Piccini JP, Hammill BG, Sinner MF, Jensen PN, Hernandez AF, Heckbert SR, Benjamin EJ, Curtis LH. Incidence and prevalence of atrial fibrillation and associated mortality among Medicare beneficiaries, 1993-2007. Circ Cardiovasc Qual Outcomes. 2012 Jan;5(1):85-93. doi: 10.1161/CIRCOUTCOMES.111.962688.
22. Mukamal KJ. The effects of smoking and drinking on cardiovascular disease and risk factors. Alcohol Res Health. 2006;29(3):199-202.
23. Kokubo Y, Watanabe M, Higashiyama A, Nakao YM, Kusano K, Miyamoto Y. Development of a Basic Risk Score for Incident Atrial Fibrillation in a Japanese General Population - The Suita Study. Circ J. 2017 Oct 25;81(11):1580-1588. doi: 10.1253/circj.CJ-17-0277.
24. Huxley RR, Lopez FL, Folsom AR, Agarwal SK, Loehr LR, Soliman EZ, Maclehose R, Konety S, Alonso A. Absolute and attributable risks of atrial fibrillation in relation to optimal and borderline risk factors: the Atherosclerosis Risk in Communities (ARIC) study. Circulation. 2011 Apr 12;123(14):1501-8. doi: 10.1161/CIRCULATIONAHA.110.009035.
25. Voskoboinik A, Kalman JM, De Silva A, Nicholls T, Costello B, Nanayakkara S,

Prabhu S, Stub D, Azzopardi S, Vizi D, Wong G, Nalliah C, Sugumar H, Wong M, Kotschet E, Kaye D, Taylor AJ, Kistler PM. Alcohol Abstinence in Drinkers with Atrial Fibrillation. N Engl J Med. 2020 Jan 2;382(1):20-28. doi: 10.1056/NEJMoa1817591.

26. Lu Y, Guo Y, Lin H, Wang Z, Zheng L. Genetically determined tobacco and alcohol use and risk of atrial fibrillation. BMC Med Genomics. 2021 Mar 9;14(1):73. doi: 10.1186/s12920-021-00915-0.
27. Wańkowicz P, Nowacki P, Gołąb-Janowska M. Atrial fibrillation risk factors in patients with ischemic stroke. Arch Med Sci. 2019 Apr 5;17(1):19-24. doi: 10.5114/aoms.2019.84212.
28. Walldius G, Malmström H, Jungner I, de Faire U, Lambe M, Van Hemelrijck M, Hammar N. Cohort Profile: The AMORIS cohort. Int J Epidemiol. 2017 Aug 1;46(4):1103-1103i. doi: 10.1093/ije/dyw333.
29. Schnabel RB, Witt H, Walker J, Ludwig M, Geelhoed B, Kossack N, Schild M, Miller R, Kirchhof P. Machine learning-based identification of risk-factor signatures for undiagnosed atrial fibrillation in primary prevention and post-stroke in clinical practice. Eur Heart J Qual Care Clin Outcomes. 2022 Dec 13;9(1):16-23. doi: 10.1093/ehjqcco/qcac013.
30. Kim EJ, Lyass A, Wang N, Massaro JM, Fox CS, Benjamin EJ, Magnani JW. Relation of hypothyroidism and incident atrial fibrillation (from the Framingham Heart Study). Am Heart J. 2014 Jan;167(1):123-6. doi: 10.1016/j.ahj.2013.10.012.
31. Jaimes MC, Torrado LAA, Reyes NFS, Mackenzie JC, Mallarino JPU. Hypothyroidism is a Risk Factor for Atrial Fibrillation after Coronary Artery Bypass Graft. Braz J Cardiovasc Surg. 2017 Nov-Dec;32(6):475-480. doi: 10.21470/1678-9741-2017-0080.
32. Liao KM, Chen CY. Incidence and risk factors of atrial fibrillation in Asian COPD patients. Int J Chron Obstruct Pulmon Dis. 2017 Aug 23;12:2523-2530. doi: 10.2147/COPD.S143691.
33. Konecny T, Park JY, Somers KR, Konecny D, Orban M, Soucek F, Parker KO, Scanlon PD, Asirvatham SJ, Brady PA, Rihal CS. Relation of chronic obstructive pulmonary disease to atrial and ventricular arrhythmias. Am J Cardiol. 2014 Jul 15;114(2):272-7. doi: 10.1016/j.amjcard.2014.04.030.
34. Chung MK, Martin DO, Sprecher D, Wazni O, Kanderian A, Carnes CA, Bauer JA, Tchou PJ, Niebauer MJ, Natale A, Van Wagoner DR. C-reactive protein elevation in

patients with atrial arrhythmias: inflammatory mechanisms and persistence of atrial fibrillation. Circulation. 2001 Dec 11;104(24):2886-91. doi: 10.1161/hc4901.101760.

35. Dhungana SP, Nepal R, Ghimire R. Prevalence and Factors Associated with Atrial Fibrillation Among Patients with Rheumatic Heart Disease. J Atr Fibrillation. 2019 Dec 31;12(4):2143. doi: 10.4022/jafib.2143.
36. Negi PC, Sondhi S, Rana V, Rathoure S, Kumar R, Kolte N, Kumar R, Rao S, Diman A, Mahajan K, Dev M, Kandoria A, Ganju N, Bhardwaj R, Merwaha R, Sharma R, Asotra S. Prevalence, risk determinants and consequences of atrial fibrillation in rheumatic heart disease: 6 years hospital based-Himachal Pradesh- Rheumatic Fever/ Rheumatic Heart Disease (HP-RF/RHD) Registry. Indian Heart J. 2018 Dec;70 Suppl 3(Suppl 3):S68-S73. doi: 10.1016/j.ihj.2018.05.013.
37. Okello E, Wanzhu Z, Musoke C, Twalib A, Kakande B, Lwabi P, Wilson NB, Mondo CK, Odoi-Adome R, Freers J. Cardiovascular complications in newly diagnosed rheumatic heart disease patients at Mulago Hospital, Uganda. Cardiovasc J Afr. 2013 Apr;24(3):80-5. doi: 10.5830/CVJA-2013-004.
38. Sharma SK, Verma SH. A Clinical Evaluation of Atrial Fibrillation in Rheumatic Heart Disease. J Assoc Physicians India. 2015 Jun;63(6):22-5. PMID: 26710395.
39. Alshehri AM. Stroke in atrial fibrillation: Review of risk stratification and preventive therapy. J Family Community Med. 2019 May-Aug;26(2):92-97. doi: 10.4103/jfcm.JFCM_99_18.
40. Oladiran O, Nwosu I. Stroke risk stratification in atrial fibrillation: a review of common risk factors. J Community Hosp Intern Med Perspect. 2019 Apr 12;9(2):113-120. doi: 10.1080/20009666.2019.1593781.
41. Yaghi S, Kamel H. Stratifying Stroke Risk in Atrial Fibrillation: Beyond Clinical Risk Scores. Stroke. 2017 Oct;48(10):2665-2670. doi: 10.1161/STROKEAHA.117.017084. Epub 2017 Sep 15. Erratum in: Stroke. 2017 Dec;48(12):e368. doi: 10.1161/STR.0000000000000155.
42. Rydén L, Sacuiu S, Wetterberg H, Najar J, Guo X, Kern S, Zettergren A, Shams S, Pereira JB, Wahlund LO, Westman E, Skoog I. Atrial Fibrillation, Stroke, and Silent Cerebrovascular Disease: A Population-based MRI Study. Neurology. 2021 Oct 19;97(16):e1608-e1619. doi: 10.1212/WNL.0000000000012675.
43. Sherer JA, Huang Q, Kiel DP, Benjamin EJ, Trinquart L. Atrial Fibrillation and the Risk of Subsequent Fracture. Am J Med. 2020 Aug;133(8):954-960. doi: 10.1016/j.amjmed.2020.02.012.

44. Lopes LA, Agrawal DK. Post-Operative Atrial Fibrillation: Current Treatments and Etiologies for a Persistent Surgical Complication. J Surg Res (Houst). 2022;5(1):159-172. doi: 10.26502/jsr.10020209.
45. Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folsom AR, Shea S, Guallar E, Bluemke DA, Lima JAC. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. Circ Res. 2017 Oct 13;121(9):1092-1101. doi: 10.1161/CIRCRESAHA.117.311312.
46. Schnabel RB, Yin X, Gona P, Larson MG, Beiser AS, McManus DD, Newton-Cheh C, Lubitz SA, Magnani JW, Ellinor PT, Seshadri S, Wolf PA, Vasan RS, Benjamin EJ, Levy D. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. Lancet. 2015 Jul 11;386(9989):154-62. doi: 10.1016/S0140-6736(14)61774-8.
47. Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ, Wolf PA. Independent risk factors for atrial fibrillation in a population-based cohort. The Framingham Heart Study. JAMA. 1994 Mar 16;271(11):840-4. PMID: 8114238.
48. Wong GR, Nalliah CJ, Lee G, Voskoboinik A, Chieng D, Prabhu S, Parameswaran R, Sugumar H, Al-Kaisey A, McLellan A, Ling LH, Sanders P, Kistler PM, Kalman JM. Sex-Related Differences in Atrial Remodeling in Patients With Atrial Fibrillation: Relationship to Ablation Outcomes. Circ Arrhythm Electrophysiol. 2022 Jan;15(1):e009925. doi: 10.1161/CIRCEP.121.009925.
49. Rodríguez-Reyes H, Lara-Vaca S, Ochoa-Guzmán A, Chiquete E; Registro Mexicano de Fibrilación Auricular Study Group. Obesity Paradox and 12 Month Outcome in Patients with Atrial Fibrillation. Arch Med Res. 2021 Feb;52(2):233-239. doi: 10.1016/j.arcmed.2020.10.015.
50. Proietti M, Lane DA, Lip GY. Relation of Nonvalvular Atrial Fibrillation to Body Mass Index (from the SPORTIF Trials). Am J Cardiol. 2016 Jul 1;118(1):72-8. doi: 10.1016/j.amjcard.2016.04.013.
51. Inoue H, Kodani E, Atarashi H, Okumura K, Yamashita T, Origasa H; J-RHYTHM Registry Investigators. Impact of Body Mass Index on the Prognosis of Japanese Patients With Non-Valvular Atrial Fibrillation. Am J Cardiol. 2016 Jul 15;118(2):215-21. doi: 10.1016/j.amjcard.2016.04.036.
52. Gupta V, Munjal JS, Jhajj P, Jhajj S, Jain R. Obesity and Atrial Fibrillation: A Narrative Review. Cureus. 2022 Nov 7;14(11):e31205. doi: 10.7759/cureus.31205.
53. Zhou Z, Hu D. An epidemiological study on the prevalence of atrial fibrillation in the

Chinese population of mainland China. J Epidemiol. 2008;18(5):209-16. doi: 10.2188/jea.je2008021.

54. Sugishita K, Shiono E, Sugiyama T, Ashida T. Diabetes influences the cardiac symptoms related to atrial fibrillation. Circ J. 2003 Oct;67(10):835-8. doi: 10.1253/circj.67.835.
55. Bano A, Rodondi N, Beer JH, Moschovitis G, Kobza R, Aeschbacher S, Baretella O, Muka T, Stettler C, Franco OH, Conte G, Sticherling C, Zuern CS, Conen D, Kühne M, Osswald S, Roten L, Reichlin T; of the Swiss-Investigators. Association of Diabetes With Atrial Fibrillation Phenotype and Cardiac and Neurological Comorbidities: Insights From the Swiss-AF Study. J Am Heart Assoc. 2021 Nov 16;10(22):e021800. doi: 10.1161/JAHA.121.021800.
56. Freeman JV, Simon DN, Go AS, Spertus J, Fonarow GC, Gersh BJ, Hylek EM, Kowey PR, Mahaffey KW, Thomas LE, Chang P, Peterson ED, Piccini JP; Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF) Investigators and Patients. Association Between Atrial Fibrillation Symptoms, Quality of Life, and Patient Outcomes: Results From the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF). Circ Cardiovasc Qual Outcomes. 2015 Jul;8(4):393-402. doi: 10.1161/CIRCOUTCOMES.114.001303.
57. Reynolds MR, Lavelle T, Essebag V, Cohen DJ, Zimetbaum P. Influence of age, sex, and atrial fibrillation recurrence on quality of life outcomes in a population of patients with new-onset atrial fibrillation: the Fibrillation Registry Assessing Costs, Therapies, Adverse events and Lifestyle (FRACTAL) study. Am Heart J. 2006 Dec;152(6):1097-103. doi: 10.1016/j.ahj.2006.08.011.
58. Ruperti Repilado FJ, Doerig L, Blum S, Aeschbacher S, Krisai P, Ammann P, Erne P, Moschovitis G, di Valentino M, Shah D, Schläpfer J, Stempfel S, Kühne M, Sticherling C, Osswald S, Conen D. Prevalence and predictors of atrial fibrillation type among individuals with recent onset of atrial fibrillation. Swiss Med Wkly. 2018 Sep 23;148:w14652. doi: 10.4414/smw.2018.14652.
59. Rizzo MR, Sasso FC, Marfella R, Siniscalchi M, Paolisso P, Carbonara O, Capoluongo MC, Lascar N, Pace C, Sardu C, Passavanti B, Barbieri M, Mauro C, Paolisso G. Autonomic dysfunction is associated with brief episodes of atrial fibrillation in type 2 diabetes. J Diabetes Complications. 2015 Jan-Feb;29(1):88-92. doi: 10.1016/j.jdiacomp.2014.09.002.
60. O'ROURKE R. Asymptomatic (silent) ischemia in stable cad. Hurst's the Heart.

2001:1219-20.

61. Onundarson PT, Thorgeirsson G, Jonmundsson E, Sigfusson N, Hardarson T. Chronic atrial fibrillation--epidemiologic features and 14 year follow-up: a case control study. Eur Heart J. 1987 May;8(5):521-7. doi: 10.1093/oxfordjournals.eurheartj.a062312.
62. Luo C, Wang L, Feng C, Zhang W, Huang Z, Hao Y, Tang A, Gao X. Predictive value of coronary blood flow for future cardiovascular events in patients with atrial fibrillation. Int J Cardiol. 2014 Dec 15;177(2):545-7. doi: 10.1016/j.ijcard.2014.08.102.
63. O'Neal WT, Salahuddin T, Broughton ST, Soliman EZ. Atrial Fibrillation and Cardiovascular Outcomes in the Elderly. Pacing Clin Electrophysiol. 2016 Sep;39(9):907-13. doi: 10.1111/pace.12907.
64. Rahman F, Wang N, Yin X, Ellinor PT, Lubitz SA, LeLorier PA, McManus DD, Sullivan LM, Seshadri S, Vasan RS, Benjamin EJ, Magnani JW. Atrial flutter: Clinical risk factors and adverse outcomes in the Framingham Heart Study. Heart Rhythm. 2016 Jan;13(1):233-40. doi: 10.1016/j.hrthm.2015.07.031.
65. He W, Chu Y. Atrial fibrillation as a prognostic indicator of myocardial infarction and cardiovascular death: a systematic review and meta-analysis. Sci Rep. 2017 Jun 13;7(1):3360. doi: 10.1038/s41598-017-03653-5.
66. Aronow WS. Hypertension associated with atrial fibrillation. Ann Transl Med. 2017 Dec;5(23):457. doi: 10.21037/atm.2017.10.33.
67. Middeldorp ME, Ariyaratnam JP, Kamsani SH, Albert CM, Sanders P. Hypertension and atrial fibrillation. J Hypertens. 2022 Dec 1;40(12):2337-2352. doi: 10.1097/HJH.0000000000003278.
68. Thacker EL, McKnight B, Psaty BM, Longstreth WT Jr, Dublin S, Jensen PN, Newton KM, Smith NL, Siscovick DS, Heckbert SR. Association of body mass index, diabetes, hypertension, and blood pressure levels with risk of permanent atrial fibrillation. J Gen Intern Med. 2013 Feb;28(2):247-53. doi: 10.1007/s11606-012-2220-4.
69. Kerr CR, Humphries KH, Talajic M, Klein GJ, Connolly SJ, Green M, Boone J, Sheldon R, Dorian P, Newman D. Progression to chronic atrial fibrillation after the initial diagnosis of paroxysmal atrial fibrillation: results from the Canadian Registry of Atrial Fibrillation. Am Heart J. 2005 Mar;149(3):489-96. doi: 10.1016/j.ahj.2004.09.053.