



Optimizing anticoagulant use for stroke prevention: a systematic review of machine learning interventions

Feras Almarshad ¹, Abdurrahman Mohammad Alshahrani ¹, Abdurrahman Saad Al Faiz ¹, Aamir Abbas ², Muhammad Shabbir ¹

ABSTRACT

Objectives: To identify and evaluate the use of machine learning interventions for optimizing anticoagulant use in stroke prevention among individuals at risk of stroke.

Methods: This systematic review adhered to the PRISMA guidelines. Searches were conducted in PubMed and Google Scholar using MeSH terms such as "Machine Learning," "Artificial Intelligence," "Anticoagulants," and "Decision Support System." Out of 333 screened articles, 36 were shortlisted based on titles, 24 after abstract review, and 15 after full-text evaluation. Included articles focused on machine learning's role in optimizing anticoagulant use for stroke prevention and analyzing primary data. Data were extracted on study design, sample size, machine learning models used, and focus areas. Risk of bias was assessed using the Prediction Model Risk of Bias Assessment Tool.

Results: Machine learning models, like logistic regression, deep neural networks, random forests, and XGBoost, outperformed traditional scoring systems like CHADS2 and CHA2DS2-VASc in predicting stroke risk. These models facilitated personalized treatment plans by incorporating genetic and metabolic data for dose optimization. Studies demonstrated the potential for machine learning to improve adherence to stroke prevention strategies, optimize anticoagulant doses, and enhance the rigor of observational studies. However, limitations included reliance on observational data and challenges in external validation and clinical utility assessments.

Conclusion: Machine learning interventions show promise in optimizing anticoagulant use for stroke prevention, surpassing traditional tools in risk stratification and treatment personalization. However, further research is needed to validate these models in clinical settings and assess their impact on patient outcomes and adherence to stroke prevention strategies.

Keywords: Machine Learning (MeSH); Artificial Intelligence (MeSH); Anticoagulants (MeSH); Stroke Prevention (Non-MeSH); Atrial Fibrillation (MeSH); Risk Stratification (Non-MeSH); Intelligence (MeSH).

THIS ARTICLE MAY BE CITED AS: Almarshad F, Alshahrani AM, Faiz ASA, Abbas A, Shabbir M. Optimizing anticoagulant use for stroke prevention: a systematic review of machine learning interventions. *Khyber Med Univ J* 2024;16(4):334-41. <https://doi.org/10.35845/kmu.2024.23682>

- 1: Department of Medicine, College of Medicine, Shaqra University, Saudi Arabia
- 2: Department of Global Health, PHC Global, Pakistan

Email [✉]: aamir.abp@gmail.com
Contact #: +92-332-2654601

Date Submitted: April 28, 2024
Date Revised: December 08, 2024
Date Accepted: December 16, 2024

and increased computational capacity has created tremendous opportunities for using machine learning techniques to solve problems related to healthcare.⁵ Machine learning has the ability to process a large amount of data and detect patterns within the data to facilitate risk stratification, prediction, and decision optimization, ensuring that most of the contextually relevant evidence for individual patients is used in real-time.^{6,7}

Though machine learning has the ability to optimize the use of anti-coagulants to prevent stroke among individuals who are at risk of stroke, the evidence in this area is scattered and, therefore, is required to be organized to understand different available machine learning techniques that can be utilized for this purpose. It is also important to check whether the ability of these techniques to improve morbidity and mortality has been checked through different interventional studies or not. Therefore, with this background, we conducted a systematic review to identify machine learning interventions for optimizing the use of anti-coagulants for stroke prevention among individuals who are at risk of stroke.

METHODS

For the purpose of this systematic review to ensure methodological rigor we adhered with Preferred Reporting Items for Systematic Reviews (PRISMA) checklist.⁸

Inclusion and exclusion criteria: We included all the articles that gave some

INTRODUCTION

Stroke is an important cause of morbidity and mortality around the world, as globally, it causes 5.5 million deaths annually.¹ Additionally, it has also been reported that around 50% of stroke survivors remain chronically disabled.¹ Only in the United States around 700,000 individuals get stroke, contributing to 165,000 deaths on an annual basis.² The economic burden of stroke ranges from USD 1,810 to 325,108.³ Stroke risk is increased by conditions like atrial fibrillation and

large-atherosclerotic disease. Anti-coagulants are used to avert the risk of stroke in such around half of these individuals.⁴ An important challenge occurs when clinicians have to choose the right type of anti-coagulant and select an appropriate dose for that anti-coagulant. Therefore, it is important to optimize the choice of these anti-coagulants based on several factors specific to the patient, including the relevant comorbidities.

In recent years, the improvement in the ability to gather large amounts of data

insights into the role of machine learning in optimizing the use of anti-coagulants in the prevention of stroke. Therefore, any article that suggested machine learning as a possible intervention to prevent stroke through optimizing the dose of anticoagulants in among individuals who are at increased risk of stroke was included irrespective of their study design. We restricted our search to articles that were published in English. Additionally, we restricted our search to articles that did analysis on primary data.

Search strategy: Relevant articles published till March 2024 were identified and included for further screening in this systematic review. We did not restrict our search to a start date to ensure inclusion of the maximum possible relevant article in this systematic review. The selected articles were screened and reviewed by two independent reviewers. The articles check for the relevance of the study inclusion criteria based on the title, abstract and the complete article and figure 1 shows the details of the articles that were included in this study. Literature was identified using Different MeSH terms that included Machine Learning, Artificial Intelligence, Anticoagulants, and Decision Support

System were used. Different variants of these terms were used by combining these MeSH terms and the variants of these search terms with appropriate Boolean operators as can be seen in Appendix 1. We identified relevant articles from PubMed and Google Scholar. Further, to ensure that we do not miss any relevant articles, we searched the reference list of the individual articles. The included articles were identified and then screened for eligibility to be included in the study. We identified a total of 333 articles, upon reading the titles of these articles, we identified 25 articles based on abstracts, and upon reading the full articles, we included 15 articles in our systematic review upon assessment of the full articles for the eligibility criteria.

Data extraction: The extracted data was entered into an excel sheet containing the information about the authors, date of publication, sample size of the study and the type of machine learning model that was used in the study. Additionally, the studies were classified into thematic areas for qualitative synthesis.

Risk of bias assessment: The methodological rigor of the included studies was assessed by carrying out risk

of bias assessment using prediction model risk of bias assessment tool.⁹ The studies were assessed based on the information related to adequacy of the data sources, explicit definitions and relevance of predictors, relevance of the outcome to the scope of the study, methodological rigor in identifying the outcome and quality of analysis. The results of risk assessment were summarized in the table II.

RESULTS

Risk stratification and stroke prediction: Though CHADS2-VASc score is used for risk stratification and, therefore, treatment optimization for individuals anticipated to have a higher risk for stroke due to atrial fibrillation, it was recently found that well-designed machine learning models like deep neural network, logistic regression, XGBoost and Random forest perform better in terms of prediction of risk of ischemic stroke in this population as compared to the CHADS2-VASc score.¹⁰ For this analysis based on a large Korean national dataset, a deep learning model was computed to identify 32 features that risk stratifying these individuals in terms of their risk for atrial fibrillation. This study was done on data from 754,949 patients, among which 62,226 were positive for ischemic stroke. Such a large dataset with detailed information about several relevant features is a unique strength that can be made possible through effective data integration and formulation of data governance and data processing pipelines at the level of institutions and countries. This can help formulate contextually relevant machine learning models that can identify individuals at high risk of stroke and, therefore, suggest anticoagulation therapy. Additionally, future studies would be required to optimize the type of anticoagulation therapy as well as their doses in the context of the contextually relevant factors identified through these machine learning models.

One study was done to check the performance of different machine learning like random forest models to predict prognosis among acute ischemic stroke patients.¹¹ Along with several other factors, it was found that anti-coagulant use was related to bad

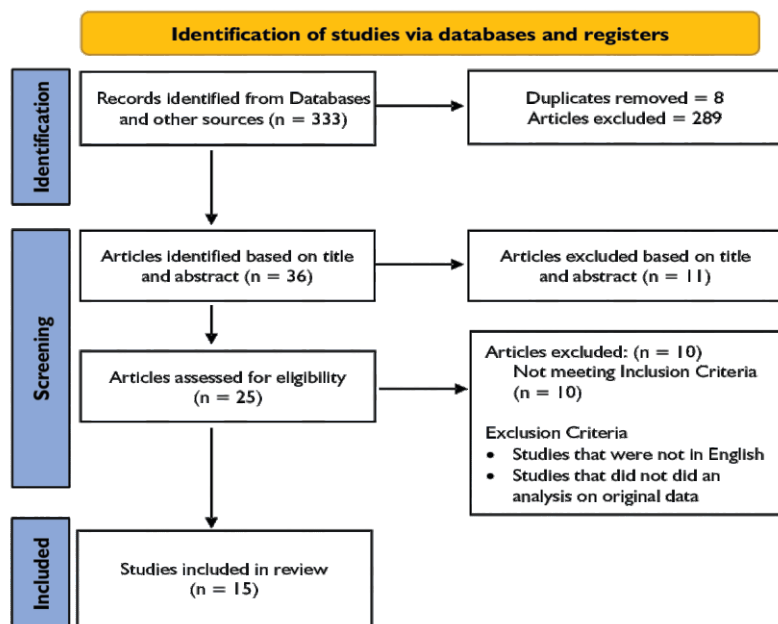


Figure 1: PRISMA Flow chart to show studies included in the study

Table I: Summary of the studies included in this systematic review

Study	Focus Area	Population	Sample Size	Machine Learning Models or other data science techniques used
Jung S, et al., (2022) ¹⁰	Prediction of ischemic stroke	Korean	754,949	Logistic Regression Attention-based deep neural network.
Kostev K, et al.,(2021) ¹³	Prediction of stroke diagnosis	German	39,652	Sub-Population Optimization and Modeling Solutions
Wang K, et al., (2024) ¹¹	Prediction of acute ischemic stroke	Chinese	677	Logistic Regression
Lip GYH, et al., (2021) ¹²	Predict of stroke risk among newly diagnosed non anticoagulated atrial fibrillation patients	US	6,457,412	Logistic Regression
Han S, et al., (2022) ²⁵	Deriving propensity score	Korean	129,434	Logistic Regression Generalized boosted model
Asiimwe, IG, et al., (2022) ¹⁷	Warfarin dose optimization	Uganda	270	Bagged trees Cubist Random Forest Regression Bayesian Additive Regression trees Model trees eXtreme Gradient Boosting Boosted trees Quantile Regression Nonlinear Least Square Regression Robust Regression Support Vector Regression Ridge Regression Partial Least Square Regression Ordinary Least Square Regression Principal Component Regression LASSO Regression Elastic Net Regression Artificial Neural Networks Multivariate Adaptive Regression Splines Regression Trees K nearest neighbors
Jahmunah, et al., (2023) ¹⁸	Warfarin dose prediction	US	6000	Deep Neural Network
Ma Z, et al., (2018) ¹⁹	Warfarin dose prediction	US	5,743	Ensemble of Machine Learning Algorithms
Kozieł-Siołkowska et al., (2022) ²³	Prediction of adherence to stroke prevention strategies	Balkan	2,712	Random Forest
Huang YC, et al., (2022) ¹⁴	Prediction of vascular events	Taiwan	11,803	Logistic Regression Naïve Bayes Random Forest Regression Tree Extreme gradient boosting
Elkin PL, et al., (2024) ¹⁵	Identify nonvalvular atrial fibrillation	US	96,681	Natural Language Processing
Huang YC et al., (2023) ²¹	Risk evaluation of patients of non-valvular atrial fibrillation taking dabigatran	Taiwan	18,113	Logistic Regression Naiva Bayes Decision Tree Random Forest XGBoost
Nadarajah R, et al., (2023) ¹⁶	Identification of patients with Atrial Fibrillation	UK	1,955	Machine Learning based Clinical Decision Support System
Meid AD, et al., (2022) ²²	Individual Treatment Effects	German	29,901	Random Forst

prognosis of individuals who suffered from acute ischemic stroke. Such findings suggesting an adverse effect of anti-coagulants among stroke patients are required to be interpreted with caution. It is possible that anti-coagulants used in acute ischemic stroke patients may increase the risk of bleeding among these individuals, therefore contributing to the worse prognosis of these patients. It is, therefore, very important to carefully decide the doses of anti-coagulants to prevent stroke among individuals who are at risk of stroke. Though this study just captured the information about whether anti-coagulants were taken or not taken but it is important to get information about the types of anti-coagulants and the doses of various types of anti-coagulants which were used in this study. Therefore, it is important that such studies be designed by integrating data science principles with epidemiological acumen to comment on the question of whether the relationship between anti-coagulants and adverse prognosis reported among patients of acute ischemic stroke was of a causal nature or it was just an association. Additionally, which type of anti-coagulant has been reported to have a better prognosis as compared to the others is also required to be explored through future studies which may compare the prognosis of acute ischemic patients between different types of anti-coagulants.

In another study, it was found that among individuals who have newly diagnosed atrial fibrillation but not receiving anti-coagulants, it was found the prediction of stroke based on different machine learning models outperformed prediction based on different algorithms like CHADS2 and CHA2DS2-VASc.¹² This study showed traditional algorithms like CHADS2 and CHA2DS2-VASc, which are regularly used in clinical practice, are required to be compared with machine learning algorithms in which a large amount of factors, including multiple comorbidities, are used in the prediction of stroke. Like several other studies done to evaluate the performance of machine learning models, this study was also an

observational study; therefore, it is also the need of the time to check the ability of these machine learning models as potential interventions to improve the ability to risk stratification and therefore, prevention of stroke in the real world.

Another study that was done in Germany intended to risk stratify through the prediction of ischemic stroke among patients who are treated with non-vitamin K antagonists oral anti-coagulants through different machine learning models.¹³ Rather than only relying on the predictions of machine learning models, this study tested the performance of subpopulation optimization and modeling solutions, which was the tool used to identify subgroups that are at an increased risk of stroke. The observational nature of the data used in this study remains an important limitation.

A secondary analysis of a randomized controlled trial dataset was done to identify the risk factors of getting vascular events using machine learning models among individuals with nonvalvular atrial fibrillation taking dabigatran.¹⁴ It was found that among many other factors, the dosage of dabigatran was an important factor that affected the adverse outcome and, therefore, is required to be carefully chosen and optimized. Random forest and XGBoost were the models that showed better predictive ability than logistic Regression. Since most of the medical records are present in the form of unstructured data, therefore it is possible to process the unstructured data through natural language processing techniques and use these data to identify individuals having nonvalvular atrial fibrillation and therefore, prevent stroke.¹⁵ Additionally machine learning models can also be used to identify undiagnosed cases of atrial fibrillation and therefore, helping in risk stratification and therefore, an opportunity to prevent stroke.¹⁶

Dose optimization: One study was done in Uganda to compare the ability of different machine learning algorithms to optimize warfarin dose.¹⁷ This study reported that simple machine learning algorithms like ordinary least square

Regression showed performance similar to more complex machine learning models. There is a need to replicate this process to use machine learning models to optimize the dosage of other oral anti-coagulants for stroke prevention. Another study used deep neural network to automate the dose prediction of warfarin based on clinical and pharmacogenetic parameters.¹⁸ Though this study captured diverse ethnic populations at the same time, it reported several challenges: the high error rate in prediction, lack of external validation of the model and lack of an assessment to assess the clinical utility of the model. Another analysis was done to improve the performance of machine learning model by using an ensemble approach to capitalize on the strengths of different machine learning models.¹⁹ In this analysis, the stacked machine learning models outperformed the linear regression models. Another study reported that it is also important to incorporate pharmacometabolomic and genetic factors while predicting the doses of warfarin.²⁰ Using the different machine learning models it is therefore, possible to plan a personalized treatment approaches to reduce risk of stroke among patients of non-valvular atrial fibrillation taking dabigatran.²¹ One study reported that individual treatment effects estimation using machine learning models can support decision-making in situations where there is inconclusive evidence.²²

Adherence to stroke prevention strategies:

One secondary analysis explored the possible predictors of adherence to stroke prevention strategies among individuals who already had atrial fibrillation.²³ This study reported that individuals who had conditions like sleep apnea, chronic kidney dialysis, and hypertrophic cardiomyopathy did not adhere to the preventive strategies. For such patients, it is required to explore the causes of non-adherence through a participatory approach taking the patient's and caregiver's perspective into account and then designing interventions to increase the probability of adhering to the recommended preventive like the use of anti-coagulants. One interventional study which tested whether the use of artificial intelligence platforms in mobile

Table II: Risk of bias assessment

Study	Participants	Predictors	Outcome	Analysis	Overall
Jung S, et al., (2022) ¹⁰	+	+	+	+	+
Kostev K, et al., (2021) ¹³	+	+	+	+	+
Wang K, et al., (2024) ¹¹	+	-	+	+	+
Lip GYH, et al., (2021) ¹²	+	-	+	+	+
Han S, et al., (2022) ²⁵	+	-	+	+	+
Asiimwe, IG, et al., (2022) ¹⁷	+	+	+	+	+
Jahmunah, et al., (2023) ¹⁸	+	+	+	+	+
Ma Z, et al., (2018) ¹⁹	-	-	+	+	-
Kozielec-Siołkowska M, et al., (2022) ²³	+	+	+	+	+
Labovitz DL, et al., (2022) ²⁴	+	-	+	+	+
Huang YC, et al., (2022) ¹⁴	+	+	+	+	+
Elkin PL, et al., (2024) ¹⁵	+	-	+	+	+
Huang YC, et al., (2023) ²¹	+	+	+	+	+
Nadarajah, R, et al. (2023) ¹⁶	+	-	+	+	-
Meid AD, et al., (2022) ²²	+	+	+	+	+

phones improves the adherence to among patients who are on anti-coagulant therapy.²⁴ It was found that as compared to the control group, the group that used the artificial intelligence platform was reported to have improved adherence to oral anti-coagulants.

Improving rigor of observational studies: Different machine learning models can be used to improve the methodological rigor of studies performed to check the comparative effectiveness of using oral anti-coagulants among patients having atrial fibrillation.²⁵ One such study was done to compute propensity scores using machine learning in such effectiveness studies and, therefore, contributing to the methodological rigor of these studies. In this study, it was reported that propensity scores computed based on generalized boosted models performed better in terms of computing propensity scores that reduced residual confounding as compared to logistic regression. Analysis in this study could have been expanded by comparing the reduction in residual confounding using other

models apart from generalized boosted models and residual confounding.

DISCUSSION

This systematic review highlights the potential of machine learning algorithms to enhance decision-making in anticoagulant therapy for stroke prevention.

We found that several machine learning algorithms are identified in different studies that can help suggest the appropriate anti-coagulants and their doses. One important finding is that some of these machine learning algorithms performed better than commonly known and used scores like CHADS2 and CHADS2-VASc.²⁶ Though this is a tempting finding to adapt to the clinical practice, it is important that the machine learning models that replace CHADS2 and CHADS2-VASc optimize the choice of different anti-coagulant therapies and their doses, are made on the datasets on which datasets having representation from the population. This brings into the discussion the question of data representativeness as model performance may differ based on

demographic, comorbidities, genetic predispositions and several other factors.²⁷ Real world data is complex, characterized by heterogeneity and there is variation in data quality and there are issues related to missing data. It is therefore important that for formulating such machine learning models it is important to formulate data flow pipelines which are a part of a larger integrated data governance framework which standardizes data collection processing protocols.²⁸ Additionally, data coming from several sources maybe useful and therefore, data partnership frameworks formulated to combine data from different sources that include electronic health records, claims databases, and patient registries may improve the generalizability of such machine learning algorithms.

Machine learning algorithms hold potential for optimizing anticoagulant dosage in stroke prevention by integrating pharmacometabolomic and pharmacogenetic data into clinical care. However, this integration presents significant challenges in clinical settings, including the need for robust data infrastructures and clinicians' ability to interpret complex datasets effectively. Moreover, the external validity of these models remains problematic, as they are typically tested on the same datasets used for training, raising concerns about their generalizability. An interventional study is required to determine the practical utility of these models.²⁹ The Use of a machine learning model for improving the quality of the evidence generated from observational studies related to the use of anti-coagulants for stroke prevention is an interesting area where individuals working on improving innovative research methods need to focus on this area.

The ethical considerations surrounding the use of machine learning models in stroke prevention are critical. Key issues such as patient privacy, informed consent, and potential algorithmic biases must be carefully addressed during data collection, model development, and implementation.³⁰ Addressing these concerns is essential for ensuring transparency, fostering trust among patients and healthcare providers, and promoting the

responsible integration of these algorithms into clinical care.³¹ Additionally, the cost-effectiveness of machine learning models must be evaluated, balancing their benefits against financial investments. Ensuring an optimal cost-benefit tradeoff will enhance the practicality and sustainability of applying these models in real-world clinical settings.

Despite the opportunities presented by machine learning models in clinical care, their integration remains challenging. To ensure accessibility, these models must be implemented through user-friendly applications that allow healthcare providers with minimal or no coding expertise to incorporate them into clinical decision support systems. Developing such applications requires an interdisciplinary approach that combines evidence from existing research with contextually relevant qualitative and quantitative data.

Furthermore, it is essential to design interpretable algorithms that enable clinicians to understand the rationale behind decisions made by these models. This transparency can enhance trust and confidence among both healthcare providers and patients, ultimately supporting the adoption of machine learning in clinical decision-making.

CONCLUSION

Machine learning algorithms show significant potential in optimizing anticoagulant use for stroke prevention, offering superior performance in risk stratification and treatment personalization compared to traditional tools like CHADS2 and CHA2DS2-VASc. These algorithms can aid in dose optimization and improve stroke prognosis predictions among at-risk populations. However, their integration into clinical practice requires careful interpretation and validation. Future research should focus on gathering contextually relevant quantitative and qualitative data to refine these models, assess their clinical utility, and evaluate their impact on patient outcomes and adherence to stroke prevention strategies.

REFERENCES

1. Donkor ES. Stroke in the 21st

Century: A Snapshot of the Burden, Epidemiology, and Quality of Life. *Stroke Res Treat* 2018; 2018:3238165. <https://doi.org/10.1155/2018/3238165>

2. Ingall T. Stroke--incidence, mortality, morbidity and risk. *J Insur Med* 2004;36(2):143-52.

3. Rochmah TN, Rahmawati IT, Dahlui M, Budiarto W, Bilqis N. Economic Burden of Stroke Disease: A Systematic Review. *Int J Environ Res Public Health* 2021;18(14):7552. <https://doi.org/10.3390/ijerph18147552>

4. Grimaldi-Bensouda L, Le Heuzey JY, Ferrières J, Leys D, Davy JM, Martinez M, et al. Stroke Prevention by Anticoagulants in Daily Practice Depending on Atrial Fibrillation Pattern and Clinical Risk Factors. *Stroke* 2021;52(10):3121-31. <https://doi.org/10.1161/STROKEAHA.120.032704>.

5. Jayatilake SMDAC, Ganegoda GU. Involvement of Machine Learning Tools in Healthcare Decision Making. *J Healthcare Eng* 2021; 2021:6679512. <https://doi.org/10.1155/2021/6679512>

6. A Powerful Paradigm for Cardiovascular Risk Stratification Using Multiclass, Multi-Label, and Ensemble-Based Machine Learning Paradigms: A Narrative Review - PMC. Accessed on: March 2, 2024. Available from URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8947682/>

7. Rodrigues PM, Madeiro JP, Marques JAL. Enhancing Health and Public Health through Machine Learning: Decision Support for Smarter Choices. *Bioengineering (Basel)* 2023;10(7):792. <https://doi.org/10.3390/bioengineering10070792>

8. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372:n71. <https://doi.org/10.1136/bmj.n71>

9. Wolff RF, Moons KGM, Riley RD,

Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med* 2019; 170(1):51-8. <https://doi.org/10.7326/M18-1376>.

10. Jung S, Song MK, Lee E, Bae S, Kim YY, Lee D, et al. Predicting Ischemic Stroke in Patients with Atrial Fibrillation Using Machine Learning. *Front Biosci (Landmark Ed)* 2022;27(3):80. <https://doi.org/10.31083/j.fbi2703080>.

11. Wang K, Hong T, Liu W, Xu C, Yin C, Liu H, et al. Development and validation of a machine learning-based prognostic risk stratification model for acute ischemic stroke. *Sci Rep* 2023; 13:13782. <https://doi.org/10.1038/s41598-023-40687-3>.

12. Lip GYH, Tran G, Genaidy A, Marroquin P, Estes C, Landsheft J. Improving dynamic stroke risk prediction in non-anticoagulated patients with and without atrial fibrillation: comparing common clinical risk scores and machine learning algorithms. *Eur Heart J Qual Care Clin Outcomes* 2021; 8(5):548-56. <https://doi.org/10.1093/ehjqcco/qcab037>

13. Kostev K, Wu T, Wang Y, Chaudhuri K, Reeve R, Tanislav C. Predicting the Risk of Ischemic Stroke in Patients Treated with Novel Oral Anticoagulants: A Machine Learning Approach. *Neuroepidemiology* 2021; 55(5):387-92. <https://doi.org/10.1159/000517512>

14. Huang YC, Cheng YC, Jhou MJ, Chen M, Lu CJ. Important Risk Factors in Patients with Nonvalvular Atrial Fibrillation Taking Dabigatran Using Integrated Machine Learning Scheme—A Post Hoc Analysis. *J Pers Med* 2022;12(5):756. <https://doi.org/10.3390/jpm12050756>

15. Elkin PL, Mullin S, Mardekian J, Crouner C, Sakilay S, Sinha S, et al. Using Artificial Intelligence With Natural Language Processing to Combine Electronic Health Record's Structured and Free Text

Data to Identify Nonvalvular Atrial Fibrillation to Decrease Strokes and Death: Evaluation and Case-Control Study. *JMIR Med Inform* 2021;9(11):e28946. <https://doi.org/10.2196/28946>.

16. Nadarajah R, Wahab A, Reynolds C, Raveendra K, Askham D, Dawson R, et al. Future Innovations in Novel Detection for Atrial Fibrillation (FIND-AF): pilot study of an electronic health record machine learning algorithm-guided intervention to identify undiagnosed atrial fibrillation. *Open Heart* 2023;10(2):e002447. <https://doi.org/10.1136/openhrt-2023-002447>.
17. Asimwe IG, Blockman M, Cohen K, Cupido C, Hutchinson C, Jacobson B, et al. Stable warfarin dose prediction in sub-Saharan African patients: A machine-learning approach and external validation of a clinical dose-initiation algorithm. *CPT Pharmacometrics Syst Pharmacol* 2022;11(1):60-8. <https://doi.org/10.1002/psp4.12740>.
18. Jahmunah V, Chen S, Oh SL, Acharya UR, Chowbay B. Automated warfarin dose prediction for Asian, American, and Caucasian populations using a deep neural network. *Comput Biol Med* 2023;153:106548. <https://doi.org/10.1016/j.combiomed.2023.106548>
19. Ma Z, Wang P, Gao Z, Wang R, Khalighi K. Ensemble of machine learning algorithms using the stacked generalization approach to estimate the warfarin dose. *PLoS One* 2018;13(10):e0205872. <https://doi.org/10.1371/journal.pone.0205872>
20. Huang Q, Cao L, Luo N, Qian H, Wei M, Xue L, et al. Predicting Range of Initial Warfarin Dose Based on Pharmacometabolomic and Genetic Inputs. *Clin Pharmacol Ther* 2022;111(1):257-67. <https://doi.org/10.1002/cpt.2407>.
21. Huang YC, Cheng YC, Jhou MJ, Chen M, Lu CJ. Integrated Machine Learning Decision Tree Model for Risk Evaluation in Patients with Non-Valvular Atrial Fibrillation When Taking Different Doses of Dabigatran. *Int J Environ Res Public Health* 2023;20(3):2359. <https://doi.org/10.3390/ijerph20032359>
22. Meid AD, Wirbka L, Groll A, Haefeli WE. Can Machine Learning from Real-World Data Support Drug Treatment Decisions? A Prediction Modeling Case for Direct Oral Anticoagulants. *Med Decis Making* 2022;42(5):587-98. <https://doi.org/10.1177/0272989X211064604>
23. Kozieł-Siołkowska M, Siołkowski S, Mihajlovic M, Lip GYH, Potpara TS. Predictors of Adherence to Stroke Prevention in the BALKAN-AF Study: A Machine-Learning Approach. *TH Open* 2022;6(3):e283-e290. <https://doi.org/10.1055/s-0042-1755617>
24. Labovitz DL, Shafner L, Reyes Gil M, Virmani D, Hanina A. Using Artificial Intelligence to Reduce the Risk of Nonadherence in Patients on Anticoagulation Therapy. *Stroke* 2017;48(5):1416-9. <https://doi.org/10.1161/STROKEAHA.116.016281>
25. Han S, Suh HS. Impact of Integrating Machine Learning in Comparative Effectiveness Research of Oral Anticoagulants in Patients with Atrial Fibrillation. *Int J Environ Res Public Health* 2022;19(19):12916. <https://doi.org/10.3390/ijerph191912916>
26. Zhu WG, Xiong QM, Hong K. Meta-analysis of CHADS2 versus CHA2DS2-VASc for predicting stroke and thromboembolism in atrial fibrillation patients independent of anticoagulation. *Tex Heart Inst J* 2015;42(1):6-15. <https://doi.org/10.14503/THIJ-14-4353>
27. Wang Y, Zhao Y, Therneau TM, Atkinson EJ, Tafti AP, Zhang N, et al. Unsupervised machine learning for the discovery of latent disease clusters and patient subgroups using electronic health records. *J Biomed Inform* 2019;102:103364. <https://doi.org/10.1016/j.jbi.2019.103364>.
28. Bozkurt S, Cahan EM, Seneviratne MG, Sun R, Lossio-Ventura JA, Ioannidis JPA, et al. Reporting of demographic data and representativeness in machine learning models using electronic health records. *J Am Med Inform Assoc* 2020;27(12):1878-1884. <https://doi.org/10.1093/jamia/ocaa164>.
29. Hong JC, Patel P, Eclow NCW, Stephens SJ, Mowery YM, Tenenbaum JD, et al. Healthcare provider evaluation of machine learning-directed care: reactions to deployment on a randomised controlled study. *NPJ Digit Med* 2023;6:25. <https://doi.org/10.1038/s41746-023-00766-3>
30. Martinez-Martin N, Luo Z, Kaushal A, Adeli E, Haque A, Kelly SS, et al. Ethical issues in using ambient intelligence in health-care settings. *Lancet Digit Health* 2021;3(2):e115-e123.
31. Lorenzini G, Shaw DM, Ossa LA, Elger BS. Machine learning applications in healthcare and the role of informed consent: ethical and practical considerations. *Clin Ethics* 2022;18(4). <https://doi.org/10.1177/1477509221094476>.

Appendix I

1	"Machine Learning" OR "Artificial Intelligence" OR "Clinical decision support system" OR "Deep Learning" OR "Large Language Model"
2	"Anticoagulants" OR "Blood Thinners" OR "Drugs" OR "Medicine" OR "Warfarin" OR "Coumadin" OR "Dabigatran" OR "Pradaxa" OR "Rivaroxaban" OR "Xarelto" OR "Apixaban" OR "Eliquis" OR "Edoxaban" OR "Savaysa" OR "Lixiana" OR "Unfractionated Heparin" OR "Low Molecular Weight Heparin" OR "Enoxaparin" OR "Lovenox" OR "Dalteparin" OR "Fragmin" OR "Tinzaparin" OR "Innohep" OR "Fondaparinux" OR "Arixtra" OR "Bivalirudin" OR "Angiomax" OR "Argatroban" OR "Otamixaban" OR "Danaparoid" OR "Sulodexide"
3	"Stroke Prevention" OR "Stroke" OR "transient ischemic attack" OR "TIA" OR "Cerebral infarction" OR "Cerebrovascular disease"
4	"Interventions" OR "Strategies"
5	1 AND 2
6	1 AND 3

AUTHORS' CONTRIBUTIONS

Following authors have made substantial contributions to the manuscript as under:

FA: Conception and study design, acquisition, analysis and interpretation of data, drafting the manuscript, approval of the final version to be published

AMA & ASAF: Study design, acquisition, analysis and interpretation of data, drafting the manuscript, approval of the final version to be published

AA: Study design, acquisition, analysis and interpretation of data, drafting the manuscript, critical review, approval of the final version to be published

MS: Study design, acquisition, analysis and interpretation of data, critical review, approval of the final version to be published

Authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICT OF INTEREST

Authors declared no conflict of interest, whether financial or otherwise, that could influence the integrity, objectivity, or validity of their research work.

GRANT SUPPORT AND FINANCIAL DISCLOSURE

Authors declared no specific grant for this research from any funding agency in the public, commercial or non-profit sectors

DATA SHARING STATEMENT

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



This is an Open Access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

KMUJ web address: www.kmu.edu.pk

Email address: kmu@kmu.edu.pk