

<https://doi.org/10.35336/VA-2020-E-28-32>

NEW APPROACHES FOR PREDICTING OUTCOMES IN PATIENTS WITH ATRIAL FIBRILLATION

N.A. Novikova¹, M.Yu. Gilyarov¹, A.Yu. Suvorov², A.Yu. Kuchina¹

¹First Sechenov Moscow Medical University (Sechenov University), Moscow, ²City Hospital #4, Moscow

Aim: we aimed to assess the capabilities of “machine learning” methods in predicting remote outcomes in patients with non-valvular atrial fibrillation (AF).

Methods. From 2015 to 2016 234 patients with non-valvular AF were included in the study (median age 72 (65; 79) years; 50.0% men). During the median follow-up of 2.9 (2.7; 3.2) years 42 patients died, 9 patients had non-fatal acute cerebral circulatory disorders and 3 patients had non-fatal myocardial infarction (MI). These events in 52 subjects (22.2% from all patients included) were combined into a combined endpoint (death and a nonfatal cardiovascular accident at the stage of remote observation). The first 184 patients comprised a “training” group. The next 50 patients formed the “test” group. The following methods of «machine learning» were used in the analysis: classification trees, linear discriminant analysis, the k-nearest neighbor method, support vectors method, neural network.

Results. Long-term outcomes were influenced by age, known traditional risk factors for cardiovascular diseases, the presence of these diseases, changes in intracardiac hemodynamics and heart chambers as evaluated by echocardiography, the presence of concomitant anemia, advanced stages of chronic kidney disease, and the administration of drugs associated with a more severe cardiovascular disease progression (amiodarone, digoxin). The best prognosis was created using the model of linear discriminant analysis, the complex neural network model, and the support vector machine.

Conclusion. Modern methods aimed at prognosis estimation seem to be of importance in cardiology. These methods include big data analysis and machine learning technologies. The methods require further evaluation and confirmation, and in the future they may allow correcting cardiovascular risks, using data from real clinical practice and evidence-based medicine at the same time.

Key words: atrial fibrillation; mortality; ischemic stroke; myocardial infarction; prognosis; machine learning; discriminant analysis; complex neural network

Conflict of Interest: nothing to declare.

Translated from *Journal of Arrhythmology*. 2019;26(2): 45-50; **Original article submitted:** 16.07.2019

Corresponding author: Novikova Nina, E-mail: nanovikova@mail.ru

For citation: Novikova NA, Gilyarov MYu, Suvorov AYu, Kuchina AYu. New approaches for predicting outcomes in patients with atrial fibrillation. *Journal of Arrhythmology*. 2019;26(2): 45-50; <https://doi.org/10.35336/VA-2019-2-45-50>.

Atrial fibrillation (AF) is among the most common cardiovascular diseases [1]. The management of AF is described in Russian and foreign clinical guidelines [2-3]. Numerous methods are known for assessing the risk of cardiovascular accidents and treatment complications in AF patients (CHA₂DS₂-VASc, HASBLED, ATRIA scores) [4-6].

Most of the available risk scales are designed to assess the likelihood of complications. However, few scales affect clinical and anamnestic, demographic indicators, concomitant diseases and the therapy taken. In recent years, novel methods for predicting risks associated with the use of so-called. «machine learning» [7]. The combination of these methods allows us to quickly analyze large amounts of data to assess the risk of a specific event in a particular patient. These methods are successfully used in genetics, oncology, endocrinology [8-10].

This article will describe several examples of the use of such methods in a small study evaluating long-term outcomes in patients with a non-valve form of atrial fibrillation.

MATERIALS AND METHODS

On the basis of the University Clinical Hospital №1 of I.M. Sechenov First Moscow State Medical University (Sechenov University) for the period 2015-2016 included 234 patients with non-valvular AF, median age 72 (65; 79) years, 50.0% of men and women. The median follow-up was 2.9 (2.7; 3.2) years, during which

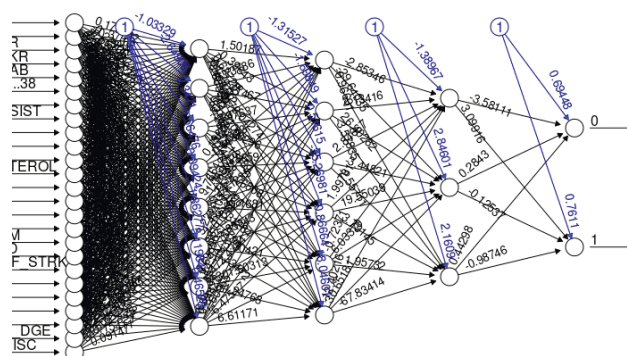


Figure 1. A multilayer perceptron with 13 “hidden” nodes used in this study.

time 42 patients died, 9 non-fatal acute cerebrovascular accidents and 3 non-fatal myocardial infarction (MI) occurred. These events in 52 patients (22.2% of all included) were summed into a combined endpoint (death and non-fatal cardiovascular catastrophes at the stage of remote observation).

The first 184 included patients formed a “training” group, the next 50 patients - the “test” group used for validation purposes. For the “training” group, regression analysis (Cox proportional risk model) was carried out to estimate the development of the combined.

Comparative analysis was carried out between the “training” and “test” groups (U - Mann-Whitney test for numerical data and Fisher’s exact test for categorical data) to make sure that there are no significant intergroup differences.

Several “machine learning” methods were used for creating various predictive models. Factors, significant in Cox regression model, were used for model training. After-

wards, the models predicted outcomes for each patient of the “test” group, thus validating the models. Because the outcomes in the “test” group were known to researchers; a comparison was made of the true results and the results predicted by the models.

The accuracy, positive and negative predictive values, Cohen kappa were calculated to assess model quality. was calculated as an independent assessment of the quality of the model. ROC analysis was carried out for a number of models.

Modeling was performed using the R language v. 3.5.1 [11].

RESULTS

The study included 234 patients with a non-valvular form of atrial fibrillation, 50% of them were men. Median age was 72 (65; 79) years. A high prevalence of cardiovascular risk factors was noted: 11.5% were smokers, 92.7% diagnosed with arterial hypertension, 24.4% had type 2

Table 1

Comparative analysis between groups, assessment of factors significantly influencing on the “training” group prognosis (Cox regression analysis)

Factor	“Training” group, (n=184)	“Test” group, (n=50)	p [#]	Relative risk	95% CI	p ^s
Age	74 (66; 79)	70 (55; 78)	0.049*	1.1	1.0-1.2	<0.001
The presence of CHF	42.90%	34.00%	0.3	2.9	1.8-4.9	<0.001
The presence of vascular diseases	39.10%	34.00%	0.6	2.5	1.5-4.2	<0.001
Prior stroke history	12.50%	18.00%	0.4	2.8	1.6-4.9	<0.001
CHA ₂ DS ₂ -VASc, points	4 (3; 5)	3 (2; 5)	0.1	1.3	1.17-1.5	<0.001
Systolic blood pressure, mm Hg	135 (120; 155)	130 (115; 150)	0.4	0.9	0.8-0.99	0.027
Duration of QRS, ms	88 (84; 96)	96 (86; 110)	0.7	1.01	1.00-1.1	0.021
Left bundle branch block	13.60%	6.00%	0.2	2.4	1.3-4.4	0.006
Prior miocardial infarction	22.80%	14.00%	0.2	2.1	1.2-3.6	0.009
Total cholesterol, mmol/l	4.4 (3.6; 5.2)	4.6 (3.8; 5.6)	0.2	0.7	0.6-0.9	0.02
Indexed volume of the LA, ml/m ²	44 (39.7; 58.2)	41.6 (35.1; 54.4)	0.07	1.02	1.01-1.1	<0.001
Indexed volume of the RA, ml/m ²	36.9 (32.4; 44.7)	36.9 (28.3; 43.3)	0.56	1.02	1.01-1.1	<0.001
SPPA, mm Hg	33 (25; 36.5)	33 (33; 40)	0.7	1.02	1.01-1.3	0.005
GFR MDRD	60 (47; 71.4)	62 (51; 75)	0.3	0.97	0.96-1.3	<0.001
CKD stage	2 (2;3)	2 (2;3)	0.4	2.1	1.6-2.6	<0.001
Digoxin at discharge, %	7.10%	10.00%	0.5	2.5	1.2-5.4	0.02
HASBLED, points	2 (2; 2)	2 (1; 3)	0.2	1.8	1.4-2.4	<0.001
History of CVD, %	52.70%	52.00%	0.9	4.2	2.3-7.7	<0.001
Stable CAD, %	37.00%	24.00%	0.09	1.7	1.1-2.8	0.04
Local hypokinesis, %	15.20%	8.00%	0.2	2.3	1.3-4.2	0.006
Global contractility decrease, %	20.70%	20.00%	0.9	2.8	1.6-4.7	<0.001
Mitral regurgitation III, %	8.20%	4.00%	0.5	3.3	1.7-6.5	<0.001
Tricuspid regurgitation III, %	7.10%	6.00%	0.9	3.3	1.6-6.7	<0.001
Anemia, %	18.50%	16.00%	0.8	3.1	1.9-5.6	<0.001
Prior amiodarone intake, %	14.10%	4.00%	0.059	2	1.1-3.7	0.03

Description: p[#] - significant differences between groups, CI - confidence interval, p^s - significant of prognosis, CAD - coronary artery disease, CHF - chronic heart failure, CKD - chronic kidney disease, CVD - cardiovascular disease, GFR - glomerular filtration rate, MDRD - Modification of Diet in Renal Disease Study, LA - left atrium, RA - right atrium, SPPA - pulmonary artery systolic pressure, * - at discharge

diabetes mellitus. Dyslipidemia was detected in 61.5% of patients.

Among the whole group, cardiovascular risks and vascular catastrophes were widely recorded. A history of coronary heart disease (CHD) was diagnosed in 26.9% of patients, while 20.9% had a history of prior MI. In 41.0% of patients, symptoms of heart failure (HF) were noted, 13.7% previously had stroke, 5.1% had transient ischemic attack. Median score on CHA₂DS₂-VASc was 4 (3; 5) points, on the HASBLED - 2 (1; 2) points.

The first group consistently included 184 patients (78.6%) in the study group, the next 50 patients (21.4%) were included in the test group. Many works show the initial ratio of the data of «training» to the «test» data of 70.0% / 30.0% to 80.0% / 20.0% [7]. Also, the separation in accordance with the sequential switching time brings the experiment somewhat closer to real practice in the form of patients who are sequentially hospitalized or who came to see a doctor.

For the “training” group, a one-dimensional Cox regression analysis was performed to identify factors that significantly affect the development of the combined endpoint during remote observation. These factors, as well as the relative risk value, are presented in Table 1:

Long-term outcomes were influenced by age, changes in intracardiac hemodynamics by echocardiography, the presence of concomitant anemia and more severe stages of CKD, as well as prescriptions associated with more severe

disease (prior amiodarone, digoxin on the long-term use at discharge).

According to these factors, a comparative analysis was carried out between the groups of «training» and «test», the results of which are given in table. 2:

Generally groups were comparable, with the exception of age (where differences were noted on the verge of significance). It is important to note that there were no significant differences either in the frequency of the combined endpoint or in the cardiovascular anamnesis.

Brief description of the methods used for “machine learning”

Classification trees

Trees combine regression and classification methods splitting of values and finding optimal threshold factor by factor. The result of numerous splitting is ultimately the classification of the object of interest. Schemes of classification trees are quite clear and understandable both by researchers and doctors. There are several algorithms for constructing classification trees (Random forest, building trees using boosting, and many others) [12].

Linear discriminant analysis

Methods underlying the linear discriminant analysis are somewhat similar to analysis of variance (ANOVA) and are associated with the search for best linear relationships between predictors. The analysis is sensitive to group size. The resulting linear interactions are somewhat reminiscent of linear regression analysis.

Table 2.

Evaluation methods predictive value classification

The method name	Accuracy	PPV	NPV	Cohen's Kappa
Classification Trees Random Forest, % (95% CI)	76.0 (61.8-86.9)	74.5 (59.7-86.1)	99.9 (37.3-99.9)	0.26
Classification tree C5.0 (with boosting), % (95% CI)	78.0 (64.4-88.5)	76.1 (61.2-87.4)	76.0 (39.8-99.9)	0.34
Simple 30-node neural network, % (95% CI)	78.0 (64.0-88.5)	80.0 (64.4-90.9)	70.0 (34.8-93.3)	0.42
Multilayer perceptron (3 layers of 17 neurons), % (95% CI)	82.0 (68.6-91.4)	79.6 (64.7-90.2)	99.9 (54.1-99.9)	0.48
Linear discriminant analysis, % (95% CI)	82.0 (68.6-91.4)	80.9 (65.9-91.3)	87.5 (47.3-99.7)	0.51
Reference vector meth, % (95% CI)	82.0 (68.6-91.4)	79.6 (64.7-90.2)	99.9 (54.1-99.9)	0.48
K-nearest-neighbours, % (95% CI)	76.0 (61.8-86.9)	74.5 (59.7-86.1)	99.9 (37.3-99.9)	0.28

Description: PPV - positive predictive value, NPV - negative predictive value, 95% CI - 95% confidence interval

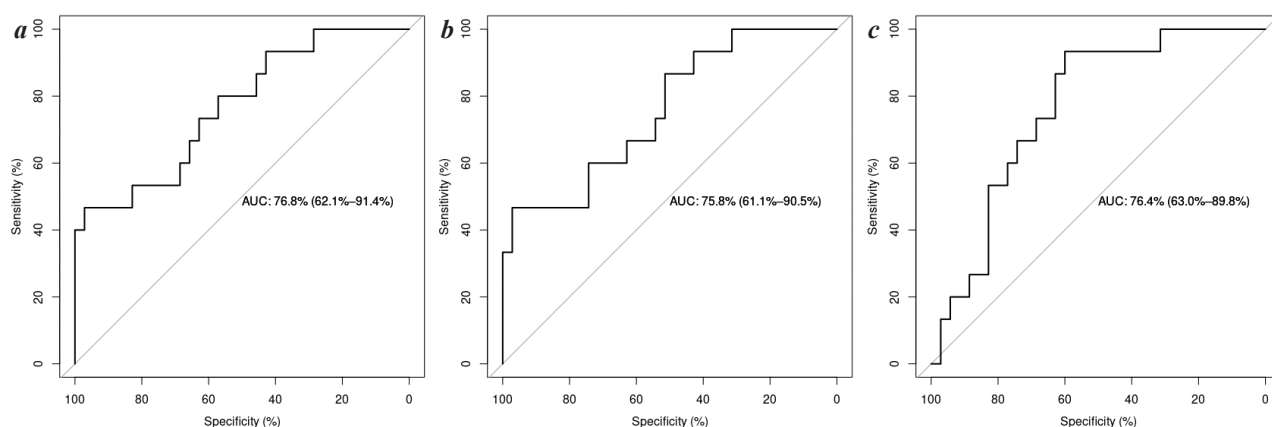


Figure 2. ROC curves for the most accurate forecasting methods in the study: a - linear discriminant analysis, b - support vector machines, c - multilayer perceptron.

The method of k-nearest neighbors

This method is close to cluster analysis when setting the number of outcome classes (in our case, there are two of them: the patient who has not reached the combined point and the patient who has reached the combined point). Each factor included in the analysis (for example, the age of the patients) is split according to the given class in terms of its average values and the distance between the average values in each class. Factors of interest are splitted thresholds are computed. Thresholds correlate with one of the predicted outcome classes.

Support vector machines

The method is associated, on the contrary, with the maximum contrasts of 2 or more classes using linear and nonlinear methods. The method aims to find the largest possible separation between classes.

Neural networks

Neural networks are based on non-linear programming algorithms. The construction of neural networks occurs in analogy with the work of neurons of the nervous system, with the presence of signals, synapses, etc. The simplest neural networks solve classification problems (simplest option is a binary classification) of an object in accordance with its characteristics. Due to the logistic functions used, the “strong” and “weak” characteristics of an object can be equalized in terms of their influence on the classification decision. The size of the neural network determine the resource consumption for its creation and operation. Usually there are “hidden» nodes structures with complex nonlinear transformations. The construction of such a network is relatively slow.

Using several layers of «hidden» nodes, allow to use nonlinear methods for more flexible and complex classification problems.

A schematic representation of the neural network used in the trial is shown in Fig. 1. Fig. 2 on the left shows the input nodes with the main patient factors, on the right - the possible predicted outcomes - “1”, when the patient is likely to develop a combined endpoint, and “0” when the patient’s state does not change. In the middle we see 3 “layers” of 6, 4 and 3 neurons. These neurons are connected to each other by the so-called coefficients - “weights” (blue and black numbers), in some way reminiscent of coefficients in regression equations. The neuron itself is a complex non-linear function that calculates the probability of a particular outcome.

Assessment of the prognostic value of classification methods

Assessment is made on the predictive accuracy of the model, positive and negative predictive value, Cohen kappa coefficient. The results are presented in table. 3:

According to the value of 95% CI, the models of linear discriminant analysis, complex neural network, and support vector machines showed the best quality. Based on the Cohen’s Kappa, the most accurate was the linear discriminant analysis model, followed by models using the support vector machines and a complex neural network.

To assess the sensitivity and specificity of the methods, ROC analysis was performed with the calculation of the area under the curve (AUC). In fig. Figure 2 shows some ROC curves of the most accurate methods:

Scientific and practical novelty

The aim of this trial was to demonstrate the possibilities of using machine learning methods in predicting long-term outcomes using typical clinical, medical, demographic characteristics of patients.

This approach will allow us to use routine examination data that the cardiologist enters in the electronic medical history, without the use of complex special scales and risk calculation techniques that require additional time from the doctor. An extremely important aspect is that the cardiologist receives information not about some abstract cardiovascular risks, but about the risk of a specific outcome or combination of outcomes in a given patient.

The scientific novelty lies in the fact that the forecasting methods are rarely used in the intersection of cardiology, epidemiology and practice. Classification prognostic models are usually built for a narrow specific problem [13]. At the same time, there can be wide perspectives for making prognosis in patients in “grey zones” of clinical guidelines. Of course, at the moment, comprehensive verification of machine learning in cardiology is needed.

DISCUSSION

Due to the fact that cardiovascular diseases are still among the three main causes of disability and mortality worldwide, the analysis of predicted outcomes is extremely important for the patient’s life, physical condition and social activity. In articles where similar forecasting methods were used, it is carefully emphasized that the prognosis should be achieved by several independent modeling methods to reduce the likelihood of false results. Factors that influenced the risk of adverse long term outcomes in this study are not unexpected (for example, more than half of those included in the history of cardiovascular accidents), and have been described in modern clinical guidelines for many years.

The analysis above showed that machine learning methods show good results with respect to prognosis. The limiting factor was a small amount of the “training” group (184 patients), the choice of endpoints (mainly the combined point without analysis of the causes of mortality). Larger sample and shorter periods can theoretically increase the accuracy of the forecast.

Currently, the abilities of large medical centers and hospitals allow aggregating a large amount of patient data, introducing electronic document management, which can serve as a “playground” for testing and implementing such methods. An important aspect is the ability of the model to provide assessment of individual risks, taking into account many factors that significantly affect the forecast.

CONCLUSION

Evaluation of long-term outcomes in patients with AF appears to be an extremely important task due to the high prevalence of the disease and the severity of complications. Modern methods aimed at predictive assessment, using big data and machine learning technologies, represent great potential for cardiology. These methods require further critical confirmation, as in the long term, they can allow correcting cardiovascular risks, using both the data of real clinical practice and the concept of evidence-based medicine.

REFERENCES

1. Chugh SS, Havmoeller R, Narayanan K et al. World-wide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation*. 2014. 25;129(8): 837-47. DOI 10.1161/CIRCULATIONAHA.113.005119.
2. Kirchhof P, Benussi S, Kotecha D et al. ESC Scientific Document Group. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J*. 2016. 7;37(38): 2893-2962. DOI 10.1093/eurheartj/ehw210.
3. Sulimov VA, Golitsyn SP, Panchenko EP, et al. Diagnosis and treatment of atrial fibrillation. *Russ J Cardiol*. 2013;18(4),S3: 5-100.
4. Camm AJ, Kirchhof P, Lip GY et al. ESC Committee for Practice Guidelines. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Europace* 2010;12(10): 1360-420. DOI 10.1093/europace/euq350.
5. Gage BF, Yan Y, Milligan PE et al. Clinical classification schemes for predicting hemorrhage: results from the National Registry of Atrial Fibrillation (NRAF). *Am Heart J*. 2006;151(3): 713-9. DOI: 10.1016/j.ahj.2005.04.017
6. Fang MC, Go AS, Chang Y et al. A new risk scheme to predict warfarin-associated hemorrhage: The ATRIA (Anticoagulation and Risk Factors in Atrial Fibrillation) Study. *J Am Coll Cardiol*. 2011; 58(4): 395-401. DOI 10.1016/j.jacc.2011.03.031.
7. Ghatak A. Machine Learning with R. *Springer Singapore*. 2017. ISBN 978-981-10-6807-2. DOI 10.1007/978-981-10-6808-9. 210 p.
8. Holder LB, Haque MM, Skinner MK. Machine learning for epigenetics and future medical applications. *Epigenetics*. 2017. 3;12(7) p.505-514. DOI 10.1080/15592294.2017.1329068.
9. Gandelman JS, Byrne MT, Mistry AM et al. Machine learning reveals chronic graft-versus-host disease phenotypes and stratifies survival after stem cell transplant for hematologic malignancies. *Haematologica*. 2019 Jan;104(1): 189-196. DOI 10.3324/haematol.2018.193441.
10. Choi BG, Rha SW, Kim SW et al. Machine Learning for the Prediction of New-Onset Diabetes Mellitus during 5-Year Follow-up in Non-Diabetic Patients with Cardiovascular Risks. *Yonsei Med J*. 2019 Feb;60(2): 191-199. DOI 10.3349/ymj.2019.60.2.191.
11. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>
12. Max Kuhn and Ross Quinlan (2018). C50: C5.0 Decision Trees and Rule-Based Models. R package version 0.1.2. URL: <https://CRAN.R-project.org/package=C50>.
13. Kim S, Kwon S, Hong SH et al. Abstract 18179: application of machine learning algorithm for early prediction to classify atrial fibrillation into paroxysmal and non-paroxysmal by analysis of fibrillatory wave on 12-lead electrocardiography. *Circulation*. 2017;136: A18179.