Challenges of Trustable AI and Added-Value on Health
B. Séroussi et al. (Eds.)

doi:10.3233/SHTI220503

The Prediction of Functional Outcome After Microsurgical Treatment of Unruptured Intracranial Aneurysm Based on Machine Learning

Timur ISHANKULOV[a], Fyodor GREBENEV[a], Uliya STRUNINA[a], Oleg SHEKHTMAN[a], Shalva ELIAVA[a] and Gleb DANILOV[a,1]
[a] *National Medical Research Center for Neurosurgery named after N.N. Burdenko, Moscow, Russian Federation*

Abstract. Our study aimed to create a machine learning model to predict patients' functional outcomes after microsurgical treatment of unruptured intracranial aneurysms (UIA). Data on 615 microsurgically treated patients with UIA were collected retrospectively from the Electronic Health Records at N.N. Burdenko Neurosurgery Center (Moscow, Russia). The dichotomized modified Rankin Scale (mRS) at the discharge was used as a target variable. Several machine learning models were utilized: a random forest upon decision trees (RF), logistic regression (LR), support vector machine (SVM). The best result with F1-score metric = 0.904 was produced by the SVM model with a label-encode method. The predictive modeling based on machine learning might be promising as a decision support tool in intracranial aneurysm surgery.

Keywords. Intracranial aneurysm, machine learning, classification, modified ranking scale, mRS

1. Introduction

Brain aneurysms occur in ~2.8% of the population and pose a severe threat to patients due to the risk of rupture and intracranial hemorrhage (1). Evidence-based management of unruptured intracranial aneurysms (UIA) might be grounded on the individual prognosis of aneurysm growth and rupture. A series of predictive modeling studies is known to address this issue. On the other hand, predicting treatment outcomes might be valued for balanced decision-making. Despite the apparent benefits of preventive treatment, surgery brings risks of disability (2.2%–10.9%) and death (0.0%–2.3%) (2). There is a lack of studies evaluating the performance of artificial intelligence in the latest task. V. Staartjes et al. (2020) demonstrated the power of machine learning to predict the outcome of unruptured aneurysm surgery in a pilot research (3). The validity of such an approach in various cohorts is a subject of rigorous investigation. Our study aimed to assess the quality of predicting the functional outcome after UIA microsurgical treatment using supervised machine learning on a single-center dataset.

---

[1] Corresponding Author, Gleb Danilov, N.N. Burdenko Neurosurgery Center, 4th Tverskaya-Yamskaya str. 16, Moscow 125047, Russian Federation; E-mail: glebda@yandex.ru.

2. Methods

Patients with unruptured intracranial aneurysms underwent microsurgery at N.N. Burdenko National Medical Research Center for Neurosurgery (Moscow, Russian Federation) in 2018-2021 were eligible for our study. We did not consider cases of consecutive UIA treatment (including endovascular) within one hospitalization, as well as aneurysms associated with other brain lesions (tumors, arteriovenous malformations, etc.). The study was approved by the Local Ethics Committee and Neurovascular Research Board at N.N. Burdenko Neurosurgery Center. Informed consent was obtained for all patients.

We collected data retrospectively from the Electronic Health Records. A set of categorical and numeric variables reflecting various patients' characteristics (age, gender, history of past illness, disease severity, neurological status, neuroimaging, intraoperative data) was first investigated via exploratory data analysis and preprocessed with feature engineering techniques. Table 1 demonstrates the final feature space we selected to use in models. Aneurysm size was evaluated via preoperative radiological data (computed tomography, magnetic resonance, or digital subtraction angiography) and matched with surgical reports and intraoperative video recordings. Thus, nine predictors related directly to patients, nine – to aneurysms, and six variables - to surgery were chosen (Table 1).

Table 1. Features related to patients, UIA and surgery exploited in predictive modeling.

	Patient's characteristics	**Aneurysm features**	**Surgical parameters**
1	Sex	Localization	Simultaneous surgery for multiple aneurysms
2	Age	Shape	Intraoperative bypass
3	American Society of Anesthesiologists (ASA) physical status	Wind neck (aneurysm neck equal or bigger than parent artery diameter)	Blood clotting disorder due to working anticoagulant/antiplatelet therapy and/or coagulopathy during the surgery
4	mRS before surgery	Size	Retrograde suction decompression or direct blood aspiration
5	History of stroke	Diverticula	
6	Number of functioning UIAs	Vessels/Nerve structures involvement	Neurosurgeon's experience (frequently or rarely operating)
7	Symptomatic type of UIA treated	Calcification/atherosclerotic lesion of aneurysm or parent artery	
8	History of anticoagulants/antiplatelet agents uptake or coagulopathy	Intraluminal thrombosis	
9	History of other intracranial aneurysms treatment	Repeated UIA treatment	

The majority of non-binary quantitative and multilevel categorical variables were collapsed into categorical to reduce the number of within-variable strata.

A modified Rankin Scale (mRS) assessment on the day of discharge was chosen as the basis for the target variable construction. mRS after the surgery varied from 0 (no symptoms) to 6 (death) and showed a pronounced imbalance which we tried to address primarily with dichotomization (Figure 1B). The binary target variable took a value of 0 if postoperative mRS was 1 or less and a value of 1 if it was equal to 2 or greater (Figure 1B).

Several machine learning models were used to predict the target variable: a random forest over decision trees (RF), logistic regression (LR), support vector machine (SVM). We also applied two labeling methods: simple label and one-hot encoding for each machine learning model. Thus, a total of 6 models were introduced.

Each test was performed after the data were randomly sampled into training (80%) and testing (20%) subsets with stratification. The model was trained on a training subset; 5-fold cross-validation (CV) was utilized to evaluate the model's quality before the final testing. Each machine learning model was tested 300 times with stratified resampling (1800 tests in total).

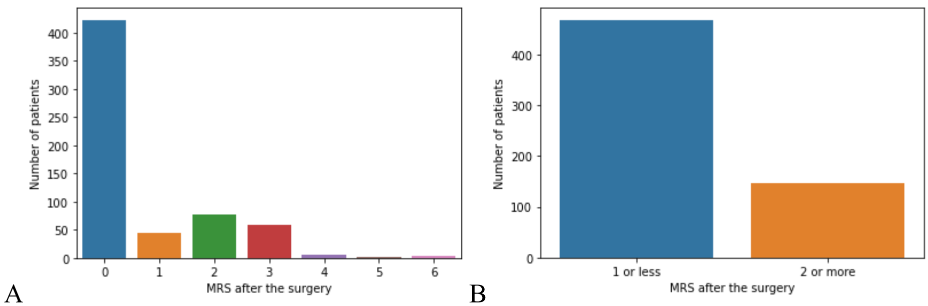

Figure 1. A - distribution of original mRS upon admission; B - distribution of binary target variable in the preprocessed dataset.

We used standard metrics to evaluate the test results: accuracy on validation samples within the cross-validation (*CV*), *Accuracy*, *Precision*, *Recall* and *F1-score* on testing samples. The results for each machine learning model were averaged across all metrics to account for the random data split and reduce the margin of error.

The exploratory data analysis was performed within the R programming environment (version 4.0.3) in RStudio Server IDE (version 1.3.1093). The modelling was done using the Python programming language (version 3.7) with the *pandas, numpy* and *sklearn* libraries in Jupyter Notebook.

3. Results

A total of 615 patients were enrolled in the study. In 31 cases, aneurysm size could not be retrieved and was denoted as "not specified". The "zeros" class has 468 points, and "ones" included 147 values (Figure 1, B).

The results of our classification experiments are presented in Table 2 in descending order of F1-score.

Table 2. UIA microsurgery outcome prediction quality with three machine learning models and two labeling methods. The 95% confidence intervals for F1 (in square brackets) were obtained via bootstrapping.

Model	Encoding	CV	Precision	Recall	Accuracy	F1-score
SVM	Label	0.922	0.951	0.878	0.925	0.904 [0.901, 0.907]
LR	Label	0.922	0.949	0.878	0.924	0.903 [0.900, 0.906]
SVM	One-hot	0.922	0.949	0.876	0.923	0.902 [0.899, 0.905]
LR	One-hot	0.919	0.941	0.872	0.918	0.896 [0.893, 0.899]
RF	One-hot	0.913	0.927	0.872	0.914	0.892 [0.888, 0.895]
RF	Label	0.911	0.926	0.871	0.913	0.891 [0.888, 0.895]

The best result in terms of the F1-score metric was 0.904, achieved with the SVM model and label-encoding method. However, the other models demonstrated close results. The results support the hypothesis that the binary outcomes of UIA surgery are pretty well separable using the proposed feature set despite imbalance.

Transformation of variables into categorical type with a reduced number of strata led to a better machine learning performance. Our resampling approach enabled calculations with a low margin of error (< 0.005). Label encoding exerted a minor influence on the results.

4. Discussion

As the availability and quality of neuroimaging improve, the number of patients diagnosed with UIA increases. In this regard, the number of operations on UIAs is also growing. For example, the number of operations for brain aneurysms in the Russian Federation increased from 1278 (in 2007) up to 7281 (in 2017), with approximately 40% of surgery performed for UIA [8]. In 60.0%-62.7% of cases, microsurgical clipping was preferable for UIA treatment in our country. A successful operation saves the patient from aneurysm rupture. On the other hand, any surgical treatment carries the risks of disability and mortality.

The most rational approach in UIA management is to compare the probability of aneurysm growth and rupture during observation with the hazards of surgical treatment. Despite the studies UCAS (4), PHASES (5,6), and ELAPSS (7,8) could give a numerical answer to the question: “What are the risks of aneurysm growth and rupture in observation?”, the UIATS study (9,10) could not answer the similar question regarding surgical interventions.

A way to cope with the prediction task is to use machine learning, which was successfully done by V. Staartjes et al. in 2020 (3). The authors demonstrated good models performance (AUC of 0.63–0.77 and the accuracy of 0.78–0.91, respectively). That work differs from our approach by less amount of input data, non-accounting of such intraoperative features as thrombextraction, retrograde suction decompression, non-inclusion of patients under 17 years of age, as well as the use of UIATS scores as one of the parameters, which could be partly subjective, and also implied not only microsurgical but endovascular treatment as well. We took that into account while collecting our database.

The limitations of our study were related to a single-center proprietary dataset, a limited number of models tested, initial sample imbalance, no additional sampling to address the imbalance, restricted feature space, no prospective model validation. Future work should be aimed at overpassing these shortcomings.

5. Conclusions

In this pilot study, good quality of functional outcome prediction after microsurgical treatment of UIA was demonstrated using traditional (shallow) machine learning methods. The predictive modeling based on machine learning might be promising as a decision support tool concerning intracranial aneurysm surgery.

References

[1] Vlak MHM, Algra A, Brandenburg R, Rinkel GJE. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. Lancet Neurol. 2011 Jul 1;10(7):626–36.

[2] Algra AM, Lindgren A, Vergouwen MDI, Greving JP, Van Der Schaaf IC, Van Doormaal TPC, et al. Procedural Clinical Complications, Case-Fatality Risks, and Risk Factors in Endovascular and Neurosurgical Treatment of Unruptured Intracranial Aneurysms: A Systematic Review and Meta-analysis. JAMA Neurol [Internet]. 2019 Mar 1 [cited 2022 Mar 28];76(3):282–93. Available from: https://pubmed.ncbi.nlm.nih.gov/30592482/

[3] Staartjes VE, Sebök M, Blum PG, Serra C, Germans MR, Krayenbühl N, et al. Development of machine learning-based preoperative predictive analytics for unruptured intracranial aneurysm surgery: a pilot study. Acta Neurochir (Wien) [Internet]. 2020 Nov 1 [cited 2022 Mar 28];162(11):2759–65. Available from: https://link.springer.com/article/10.1007/s00701-020-04355-0

[4] Morita A, Kirino T, Hashi K, Aoki N, Fukuhara S, Hashimoto N, et al. The natural course of unruptured cerebral aneurysms in a Japanese cohort. N Engl J Med [Internet]. 2012 Jun 28 [cited 2022 Mar 28];366(26):2474–82. Available from: https://pubmed.ncbi.nlm.nih.gov/22738097/

[5] Greving JP, Wermer MJH, Brown RD, Morita A, Juvela S, Yonekura M, et al. Development of the PHASES score for prediction of risk of rupture of intracranial aneurysms: a pooled analysis of six prospective cohort studies. Lancet Neurol [Internet]. 2014 Jan [cited 2022 Mar 28];13(1):59–66. Available from: https://pubmed.ncbi.nlm.nih.gov/24290159/

[6] Backes D, Vergouwen MDI, Tiel Groenestege AT, Bor ASE, Velthuis BK, Greving JP, et al. PHASES Score for Prediction of Intracranial Aneurysm Growth. Stroke [Internet]. 2015 May 20 [cited 2022 Mar 28];46(5):1221–6. Available from: https://www.ahajournals.org/doi/abs/10.1161/strokeaha.114.008198

[7] Backes D, Rinkel GJE, Greving JP, Velthuis BK, Murayama Y, Takao H, et al. ELAPSS score for prediction of risk of growth of unruptured intracranial aneurysms. Neurology [Internet]. 2017 Apr 25 [cited 2022 Mar 28];88(17):1600–6. Available from: https://pubmed.ncbi.nlm.nih.gov/28363976/

[8] van Kammen MS, Greving JP, Kuroda S, Kashiwazaki D, Morita A, Shiokawa Y, et al. External Validation of the ELAPSS Score for Prediction of Unruptured Intracranial Aneurysm Growth Risk. J Stroke [Internet]. 2019 Sep 1 [cited 2022 Mar 28];21(3):340. Available from: /pmc/articles/PMC6780020/

[9] Etminan N, Beseoglu K, Barrow DL, Bederson J, Brown RD, Connolly ES, et al. Multidisciplinary consensus on assessment of unruptured intracranial aneurysms: Proposal of an international research group. Stroke [Internet]. 2014 [cited 2022 Mar 28];45(5):1523–30. Available from: https://www.ahajournals.org/doi/abs/10.1161/strokeaha.114.004519

[10] Etminan N, Brown RD, Beseoglu K, Juvela S, Raymond J, Morita A, et al. The unruptured intracranial aneurysm treatment score: A multidisciplinary consensus. Neurology [Internet]. 2015 Sep 8 [cited 2022 Mar 28];85(10):881. Available from: /pmc/articles/PMC4560059/