



Machine Learning Models can Detect Aneurysm Rupture and Identify Clinical Features Associated with Rupture

Michael A. Silva⁴, Jay Patel³, Vasileios Kavouridis^{1,2}, Troy Gallerani^{1,2}, Andrew Beers³, Ken Chang³, Katharina V. Hoebel³, James Brown³, Alfred P. See^{1,2}, William B. Gormley^{1,2}, Mohammad Ali Aziz-Sultan^{1,2}, Jayashree Kalpathy-Cramer³, Omar Arnaout^{1,2}, Nirav J. Patel^{1,2}

■ **BACKGROUND:** Machine learning (ML) has been increasingly used in medicine and neurosurgery. We sought to determine whether ML models can distinguish ruptured from unruptured aneurysms and identify features associated with rupture.

■ **METHODS:** We performed a retrospective review of patients with intracranial aneurysms detected on vascular imaging at our institution between 2002 and 2018. The dataset was used to train 3 ML models (random forest, linear support vector machine [SVM], and radial basis function kernel SVM). Relative contributions of individual predictors were derived from the linear SVM model.

■ **RESULTS:** Complete data were available for 845 aneurysms in 615 patients. Ruptured aneurysms ($n = 309$, 37%) were larger (mean 6.51 mm vs. 5.73 mm; $P = 0.02$) and more likely to be in the posterior circulation (20% vs. 11%; $P < 0.001$) than unruptured aneurysms. Area under the receiver operating curve was 0.77 for the linear SVM, 0.78 for the radial basis function kernel SVM models, and 0.81 for the random forest model. Aneurysm location and size were the 2 features that contributed most significantly to the model. Posterior communicating artery, anterior communicating artery, and posterior inferior cerebellar artery locations were most highly associated with rupture, whereas paraclinoid and middle cerebral

artery locations had the strongest association with unruptured status.

■ **CONCLUSIONS:** ML models are capable of accurately distinguishing ruptured from unruptured aneurysms and identifying features associated with rupture. Consistent with prior studies, location and size show the strongest association with aneurysm rupture.

INTRODUCTION

Existing tools to estimate rupture risk for unruptured intracranial aneurysms are limited. The International Study of Unruptured Intracranial Aneurysms (ISUIA)¹ and other natural history studies²⁻⁵ have consistently implicated size and location. The role of other clinical risk factors (e.g., smoking,⁶ multiple aneurysms,⁴ age⁷) remains unclear, and even our understanding of the importance of size^{4,5} and location^{8,9} continues to evolve. Recent attempts to understand rupture risk have taken new approaches, such as using hemodynamic modeling, morphologic analysis, and wall inflammation detection.¹⁰⁻¹⁴ Despite numerous studies seeking to better characterize the clinical and angiographic risk factors of aneurysmal rupture, our understanding remains limited.

New tools for analyzing large datasets may offer a modern approach to understanding aneurysm rupture risk. Machine

Key words

- Aneurysm
- Aneurysm rupture
- Artificial intelligence
- Machine learning
- Subarachnoid hemorrhage

Abbreviations and Acronyms

ACOMM: Anterior communicating artery

AUC: Area under the curve

ICA: Internal carotid artery

ISUIA: International Study of Unruptured Intracranial Aneurysms

MCA: Middle cerebral artery

ML: Machine learning

PCOMM: Posterior communicating artery

PICA: Posterior inferior cerebellar artery

SVM: Support vector machine

From the Departments of ¹Neurosurgery and ²Computational Neuroscience Outcomes Center, Brigham and Women's Hospital, Boston, Massachusetts; ³Department of Radiology, Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, Massachusetts; and ⁴Department of Neurosurgery, Jackson Memorial Hospital, Miami, Florida, USA

To whom correspondence should be addressed: Nirav J. Patel, M.D.
[E-mail: npatel40@bwh.harvard.edu]

Citation: *World Neurosurg.* (2019) 131:e46-e51.

<https://doi.org/10.1016/j.wneu.2019.06.231>

Journal homepage: www.journals.elsevier.com/world-neurosurgery

Available online: www.sciencedirect.com

1878-8750/\$ - see front matter © 2019 Published by Elsevier Inc.

learning (ML) is a type of artificial intelligence that can detect associations between features of a dataset without being explicitly programmed.¹⁵⁻¹⁷ Unlike classical statistics that require a hypothesis-driven approach to answering clinical questions, the statistical learning in ML allows for the discovery of unanticipated associations. ML algorithms not only detect important relationships but can be simply and rapidly applied to new data to make predictions.¹⁷ ML has been increasingly used throughout medicine, including neurosurgery, for diagnosis, surgical planning, and outcome prediction.^{15,17,18}

Using 16 years of retrospective, single-institution data, we trained 3 unique ML models to classify rupture status and identify the clinical features most strongly associated with rupture based on basic patient data.

METHODS

Patient Population

This retrospective study was approved by the Partners institutional review board (#2015P002352). Patient data were retrospectively reviewed for patients with unruptured or ruptured intracranial

aneurysms detected on vascular imaging at Brigham and Women's Hospital between 2002 and 2018. Complete data were available for 615 patients in whom 845 aneurysms were found. Clinical features were extracted from admission notes, operative reports, and discharge documentation; radiographic data were extracted from radiology reports.

Study Variables

The models were trained using the following variables: aneurysm size, location (anterior communicating artery [ACOMM], anterior cerebral artery, basilar artery, internal carotid artery [ICA] terminus, middle cerebral artery [MCA], paraclinoid ICA, posterior communicating artery [PCOMM], posterior cerebral artery, posterior inferior cerebellar artery [PICA], superior cerebellar artery, vertebral artery), aneurysm side, patient sex, patient age, family history of intracranial aneurysm, presence of multiple aneurysms, smoking status (current or former), history of hypertension, history of diabetes, and history of hyperlipidemia. Aneurysm location, size, and multiplicity was extracted from formal radiology reports. In the case of incomplete documentation, the images were primarily reviewed.

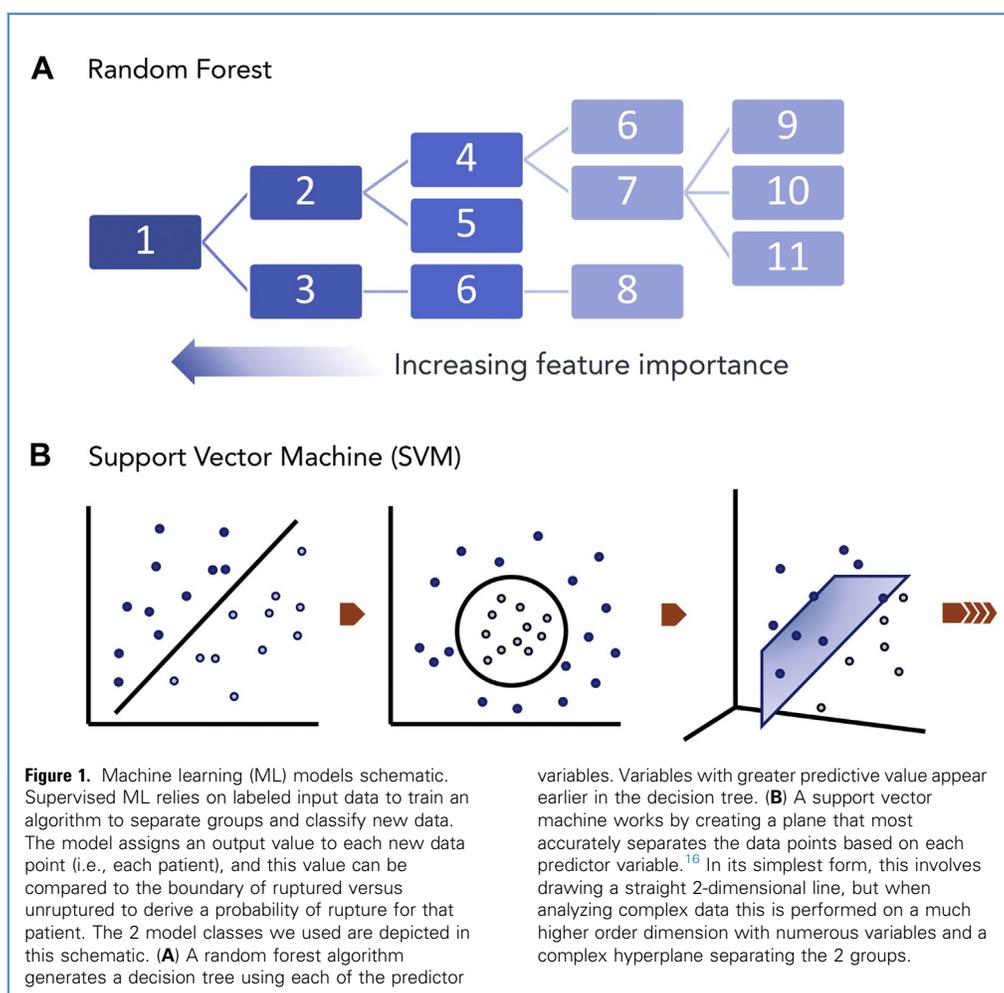


Table 1. Feature Contribution Derived from Linear Support Vector Machine Model

Feature	Feature Contribution to Linear SVM (Bootstrapped Mean)	95% CI	P Value
Paraclinoid*	-1.093	[-1.353, -0.834]	<0.0001*
MCA*	-0.822	[-1.368, -0.276]	0.003*
ICA terminus	-0.500	[-1.105, 0.106]	0.106
Right side	-0.225	[-0.488, 0.038]	0.093
Left side	-0.194	[-0.453, 0.064]	0.142
Hyperlipidemia	-0.181	[-0.517, 0.156]	0.290
Multiple aneurysms	-0.142	[-0.445, 0.162]	0.360
Smoking	-0.137	[-0.408, 0.135]	0.321
Vertebral	-0.115	[-0.278, 0.048]	0.166
Diabetes	-0.109	[-0.404, 0.186]	0.470
Family history	-0.061	[-0.319, 0.198]	0.644
PCA	0.010	[-0.278, 0.298]	0.946
ACA	0.052	[-0.551, 0.655]	0.866
Age	0.056	[-0.178, 0.290]	0.641
Hypertension	0.063	[-0.135, 0.262]	0.533
Sex	0.109	[-0.152, 0.371]	0.416
SCA	0.133	[-0.386, 0.652]	0.616
Size	0.188	[-0.129, 0.505]	0.246
Basilar	0.265	[-0.322, 0.852]	0.377
PICA*	0.654	[0.288, 1.019]	0.0004*
ACOMM*	0.660	[0.350, 0.969]	<0.0001*
PCOMM*	0.787	[0.468, 1.106]	<0.0001*

SVM, support vector machine; CI, confidence interval; MCA, middle cerebral artery; ICA, internal carotid artery; PCA, posterior cerebral artery; ACA, anterior cerebral artery; SCA, superior cerebellar artery; PICA, posterior inferior cerebellar artery; ACOMM, anterior communicating artery; PCOMM, posterior communicating artery.
*Statistically significant.

Model Training and Optimization

A random sample of 60% of the dataset was used to train the models, and 40% was reserved for testing. Training and testing sets were balanced to each include 36.5% ruptured cases. We used a random forest¹⁹ and 2 types of support vector machine (SVM)¹⁶ classifiers to determine the ability of clinically relevant features to distinguish unruptured from ruptured aneurysms. A random forest is a commonly used classifier that generates a decision tree with each branch point corresponding to a particular predictor variable (Figure 1). Much like a flowchart, the most important features appear earlier in the decision tree. A new datapoint can therefore be fed into the decision tree and classified based on each branch point. An SVM is another type of classifier that finds a hyperplane that maximally separates the 2 classes (Figure 1). In

Table 2. Study Population

Variable	Unruptured	Ruptured	P Value
Aneurysms	536 (63%)	309 (37%)	
Size (mm)	5.73 ± 4.71	6.51 ± 4.14	0.02
Anterior circulation	475 (89%)	248 (80%)	<0.001*
Multiple aneurysms	192 (36%)	57 (18%)	<0.001*
Patient age	57.4 ± 12.2	57.8 ± 13.0	0.62
Sex (Male)	96 (18%)	79 (26%)	0.008
Smoking	306 (57%)	149 (48%)	0.013
Family history	91 (17%)	37 (12%)	0.051
Hypertension	299 (56%)	173 (56%)	0.95
Diabetes	72 (13%)	34 (11%)	0.30
Hyperlipidemia	169 (32%)	69 (22%)	0.004

*Statistically significant.

other words, a line is generated to separate datapoints in a multidimensional plot. In its simplest form, this is a 2-dimensional line, but in complex multivariate analysis it assumes a higher order form. Rupture probability was calculated by determining the distance between each datapoint and the linear hyperplane. More formally, the distance to the boundary was squashed via a learned sigmoid function to output a probability value; points far from the hyperplane exhibit high probabilities of either being unruptured or ruptured and points close to the hyperplane exhibit probabilities close to 0.5 (implying uncertainty of whether the case was unruptured or ruptured). Both linear and radial basis kernel variants of SVM (as implemented in the *sklearn* package in Python [Python Software Foundation, Beaverton, Oregon, USA]) were trained and assessed for performance on the test set. Calibration plots were generated for each of the models. Binary cutoffs were chosen by maximizing the joint number of true positive and true negatives on the training set, which is equivalent to choosing the threshold that relates to the point that is closest to the (0,1) corner of the area under the curve (AUC) graph.

Classification and Feature Importance

For the model trained via linear SVM, relative feature contributions were derived from the weighted coefficients of the hyperplane. To ensure robustness of the feature contributions, we generated 1000 bootstrapped sets (in which 75% of the training set was sampled with replacement). Linear SVMs were trained on each of these bootstrapped sets. The mean, 95% confidence intervals, and P values of the hyperplane slopes (representative of individual feature contributions) were calculated and are shown in Table 1. Positive feature contribution indicates a positive association with rupture; negative feature contribution indicates a negative correlation with rupture.

Statistical Analysis

Univariate analysis to compare ruptured and unruptured groups (Tables 2 and 3) was performed with a t test for continuous

Table 3. Aneurysm Location Distribution for Unruptured and Ruptured Cohorts

Rupture Status	Total	Paraclinoid	PCOMM	ICA Terminus	MCA	ACOMM	ACA	PCA	Basilar	SCA	Vertebral	PICA
Unruptured	536	178	47	31	140	59	20	5	35	11	5	6
		33%	8.8%	5.8%	26%	11%	3.7%	1%	6.5%	2%	1%	1%
Ruptured	309	26	71	9	43	85	14	3	30	8	4	17
		8.4%	23%	3%	14%	28%	4.5%	9.7%	9.7%	2.6%	1.3%	5.5%
P value		<0.001*	<0.001*	0.06	<0.001*	<0.001*	0.57	0.96	0.095	0.61	0.62	<0.001*

PCOMM, posterior communicating artery; ICA, internal carotid artery; MCA, middle cerebral artery; ACOMM, anterior communicating artery; ACA, anterior cerebral artery; PCA, posterior cerebral artery; SCA, superior cerebellar artery; PICA, posterior inferior cerebellar artery.
*Statistically significant.

features (patient age and aneurysm size), and a χ^2 test was used for binary features. A P value cutoff of 0.05 with a Bonferroni correction yielded a significance threshold of 0.002, thus shielding against type I error in the setting of multiple comparisons. Multivariate analysis was performed using the trained SVM on the held-out test set. As previously described, feature contribution was calculated by deriving the slope of the hyperplane for each feature and bootstrapping this calculation 1000 times, and this distribution was compared against the null hypothesis of a coefficient of zero. Multivariate logistic regression was also performed on the test set. Model performance was assessed via AUC on the test set. Sensitivity, specificity, positive predictive value, and negative predictive value were also calculated. All analysis was performed in Python 3.5.

RESULTS

A total of 845 aneurysms in 615 patients were included in the study. Of the 845 aneurysms, 309 (37%) were ruptured. A total of 473 patients had a single aneurysm; 142 patients had multiple aneurysms. A summary of the patient and aneurysm characteristics for the unruptured and ruptured cohorts is provided in **Table 2**. Ruptured aneurysms were larger (mean 6.51 mm vs. 5.73 mm; $P = 0.02$) and more likely to be in the posterior circulation (20% vs. 11%; $P < 0.001$) (**Table 2**).

Aneurysm location varied significantly between the unruptured and ruptured cohorts. The paraclinoid ICA was the most common

location for unruptured aneurysms (33%), whereas just 8.4% of ruptured aneurysms were paraclinoid (**Table 3**). The ACOMM (28%) and PCOMM (23%) were the 2 most common locations for ruptured aneurysms. The groups were not significantly different in the rate of anterior cerebral artery, ICA terminus, posterior cerebral artery, basilar artery, superior cerebellar artery, and vertebral artery aneurysms (**Table 3**).

Three models were trained to differentiate between ruptured and unruptured cases: random forest, linear SVM, and radial basis function kernel SVM. A receiver operating characteristic curve was generated for each of these models to evaluate the sensitivity and specificity for classifying the rupture status of each case. The linear SVM yielded an AUC of 0.77, the radial basis function kernel SVM model had an AUC of 0.78, and the random forest model returned an AUC of 0.81 (**Table 4**). Probability of rupture for each case could then be calculated by determining the distance between each datapoint and the hyperplane generated by the SVM model. Multivariate logistic regression, for purposes of comparison, yielded an AUC of 0.79.

Further interrogation of the linear SVM model identified the predictor variables that contributed most significantly to the final model. Variables are listed and sorted by ascending contribution to the model in **Table 1**, with higher feature contribution denoting a stronger association with rupture. Negative feature contribution signifies a negative association with rupture, meaning a greater likelihood of being unruptured. By bootstrapping the model, a mean and 95% confidence interval for each feature could be generated (**Table 1**). Aneurysm location contributed more to the final model than any other feature (**Figure 2, Table 1**). The 4 variables with the most positive feature contribution and the 3 variables with the most negative feature contribution were in all locations. PCOMM, ACOMM, and PICA locations were significantly associated with rupture, whereas paraclinoid and MCA locations were significantly associated with unruptured status. After location, aneurysm size was the feature with the next most positive association with rupture.

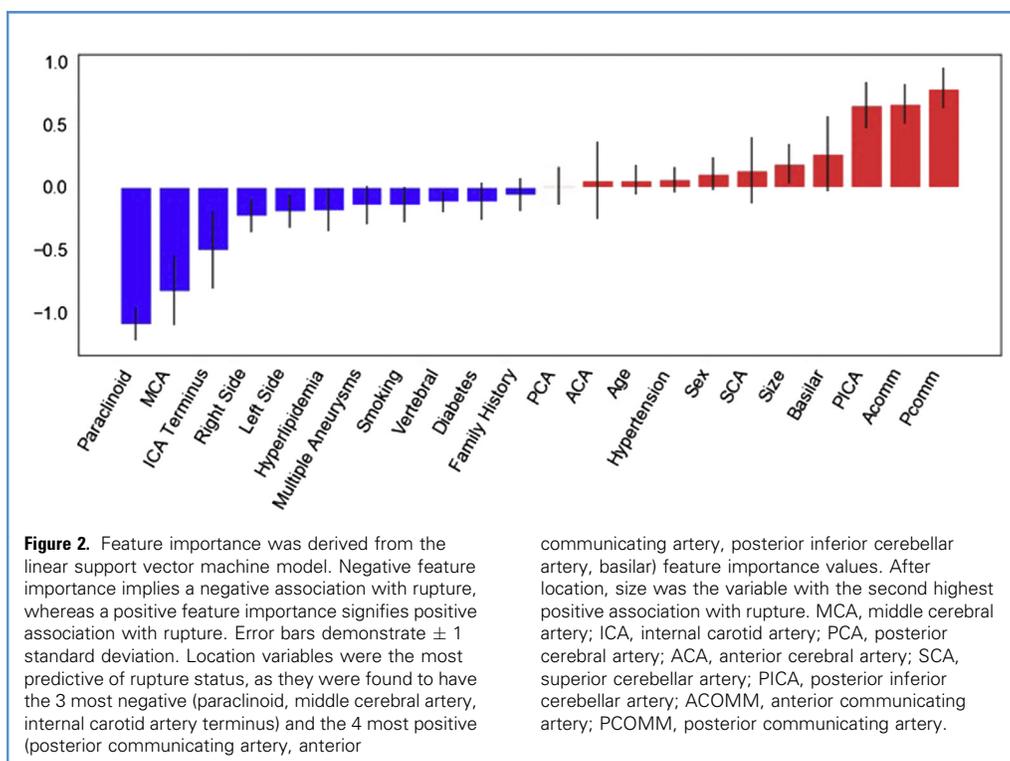
DISCUSSION

The importance of aneurysm size and location when estimating rupture risk is well-established. Our models corroborate the findings of ISUIA, unruptured cerebral aneurysms, and other natural

Table 4. Model Performance

Statistical Metric	Random Forest	Linear SVM	RBF Kernel SVM	Logistic Regression
Area under the curve	0.81	0.77	0.78	0.79
Sensitivity	0.75	0.85	0.73	0.81
Specificity	0.76	0.60	0.73	0.67
Positive Predictive Value	0.64	0.55	0.60	0.59
Negative Predictive Value	0.84	0.87	0.82	0.86

SVM, support vector machine; RBF, radial basis function.



history studies that have implicated size and location as the most significant risk factors for rupture. Moreover, the locations that our linear SVM model identified to be most positively (PCOMM, ACOMM) and negatively (paraclinoid, MCA) associated with rupture are consistent with what prior studies have demonstrated.^{1,2,7} Our model did not find smoking or multiple aneurysms to be associated with increased rupture risk. This finding is consistent with the unruptured cerebral aneurysms study, although several other studies have found them to be risk factors for rupture.^{3,4,6} Altogether, the concordance of our findings with what has been shown in the literature suggests that the methodology behind these ML approaches is valid and increasingly capable of being applied to clinical questions in neurosurgery.

A key advantage of using a linear SVM model to understand aneurysm rupture risk is the ease of quantifying the relative importance of individual risk factors. Although the importance of size and location is universally accepted, few studies comment on their relative importance. Our model suggests that location is significantly more important than size when estimating rupture risk. This might help explain the high rates of rupture for small aneurysms that has been reported by several groups.^{4,5,20} It is important to recognize that the significance of risk factors like location and size, as determined by our model, does not imply a simple, linear association. Rather, these ML models interpret each variable in the context of all other predictors, considering the complex interactions among them. The effect size of individual variables might, in fact, not be detected with classical statistics.²¹ For example, a variable with an equivalent mean value in each cohort being compared might still be significantly predictive of one class over another due to

complex interactions with other variables in the dataset. Although the SVM model provided interesting insights into the relative feature importance of predictor variables, the primary objective of this analysis was to demonstrate credibility of the technique. Indeed, the model relies on variables supported by the literature.

This study has several limitations. The generalizability of our findings is limited by the single-institution nature of our patient cohort. The patient population that made up our training dataset may bias our models such that their accuracy is diminished when tested on external data. Despite consistently high AUC values when tested on new data from our institution, each of our models must be tested on external data to assess their performance on patient data from outside our institution. Additionally, the models were trained on retrospective data comparing ruptured and unruptured cases rather than on natural history data of untreated unruptured aneurysms. Given the lack of sufficient clinical equipoise to acquire long-term follow-up data on untreated unruptured aneurysms, our strategy was to use current rupture status to train a model to interpret rupture risk. The models' confidence in classifying a case as ruptured or unruptured was treated as a proxy for rupture risk. These models are not intended to directly predict rupture risk; rather, they can distinguish ruptured from unruptured aneurysms, and they can identify variables associated with rupture. Evaluating our models on a prospective database with long-term follow-up of untreated aneurysms would help validate our findings. The addition of radiographic characteristics might also improve the accuracy of these models; in fact, geometrical parameters have been used to generate accurate ML models to predict aneurysm rupture risk.¹⁸

This study is, above all, a proof of concept demonstrating the feasibility of using ML tools for the purpose of interpreting large complex data in cerebrovascular neurosurgery. The accordance of this study's findings (i.e., the features with the greatest contribution to the model) with ISUIA and similar studies suggests that the ML methods used here are credible. With further refinement and validation on external and prospective data, the techniques described here have the potential to yield decision-making tools for neurosurgeons faced with increasing volumes of complex data. Indeed, ML has already proven useful for predicting delayed cerebral ischemia after subarachnoid hemorrhage,²² classifying brain tumors,¹⁵ predicting outcome after flow diversion,²³ and detecting aneurysms from magnetic resonance angiography.²⁴

An approach similar to the one used in our study might help address clinical questions in neurosurgery requiring interpretation of complex data.

CONCLUSIONS

The ability of ML to identify complex relationships in large datasets makes it uniquely suited to the task of aneurysm rupture classification. The models developed in this study can accurately classify rupture status using predictors that are in accordance with prior literature. Our findings suggest that the ML techniques described here are reliable and have the potential for applications throughout neurosurgery.

REFERENCES

- Wiebers DO, Whisnant JP, Huston J 3rd, et al. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. *Lancet*. 2003;362:103-110.
- UCAS Japan Investigators, Morita A, Kirino T, et al. The natural course of unruptured cerebral aneurysms in a Japanese cohort. *N Engl J Med*. 2012;366:2474-2482.
- Juvela S, Porras M, Poussa K. Natural history of unruptured intracranial aneurysms: probability of and risk factors for aneurysm rupture. *J Neurosurg*. 2000;93:379-387.
- Sonobe M, Yamazaki T, Yonekura M, Kikuchi H. Small unruptured intracranial aneurysm verification study: SUAve study, Japan. *Stroke*. 2010;41:1969-1977.
- Korja M, Lehto H, Juvela S. Lifelong rupture risk of intracranial aneurysms depends on risk factors: a prospective Finnish cohort study. *Stroke*. 2014;45:1958-1963.
- Dasenbrock HH, Rudy RF, Rosalind Lai PM, et al. Cigarette smoking and outcomes after aneurysmal subarachnoid hemorrhage: a nationwide analysis. *J Neurosurg*. 2018;129:446-457.
- Juvela S, Poussa K, Lehto H, Porras M. Natural history of unruptured intracranial aneurysms: a long-term follow-up study. *Stroke*. 2013;44:2414-2421.
- Chung BJ, Doddasomayajula R, Mut F, et al. Angioarchitectures and hemodynamic characteristics of posterior communicating artery aneurysms and their association with rupture status. *AJNR Am J Neuroradiol*. 2017;38:2111-2118.
- Doddasomayajula R, Chung BJ, Mut F, et al. Hemodynamic characteristics of ruptured and unruptured multiple aneurysms at mirror and ipsilateral locations. *AJNR Am J Neuroradiol*. 2017;38:2301-2307.
- Can A, Du R. Association of hemodynamic factors with intracranial aneurysm formation and rupture: systematic review and meta-analysis. *Neurosurgery*. 2016;78:510-520.
- Skodvin TO, Evju O, Sorteberg A, Isaksen JG. Prerupture intracranial aneurysm morphology in predicting risk of rupture: a matched case-control study. *Neurosurgery*. 2019;84:132-140.
- Rahman M, Smietana J, Hauck E, et al. Size ratio correlates with intracranial aneurysm rupture status: a prospective study. *Stroke*. 2010;41:916-920.
- Chien A, Xu M, Yokota H, Scalzo F, Morimoto E, Salamon N. Nonsphericity index and size ratio identify morphologic differences between growing and stable aneurysms in a longitudinal study of 93 cases. *AJNR Am J Neuroradiol*. 2018;39:500-506.
- Hu P, Yang Q, Wang DD, Guan SC, Zhang HQ. Wall enhancement on high-resolution magnetic resonance imaging may predict an unsteady state of an intracranial saccular aneurysm. *Neuroradiology*. 2016;58:979-985.
- Senders JT, Arnaout O, Karhade AV, et al. Natural and artificial intelligence in neurosurgery: a systematic review. *Neurosurgery*. 2018;83:181-192.
- Noble WS. What is a support vector machine? *Nat Biotechnol*. 2006;24:1565-1567.
- Senders JT, Staples PC, Karhade AV, et al. Machine learning and neurosurgical outcome prediction: a systematic review. *World Neurosurg*. 2018;109:476-486.e1.
- Aranda A, Valencia A. Study on cerebral aneurysms: rupture risk prediction using geometrical parameters and wall shear stress with CFD and machine learning tools. *Machine Learn Appl An Int J*. 2018;5.
- Breiman L. Random forests. *Machine Learn*. 2001;45:5-32.
- Bender MT, Wendt H, Monarch T, et al. Small aneurysms account for the majority and increasing percentage of aneurysmal subarachnoid hemorrhage: a 25-year, single institution study. *Neurosurgery*. 2018;83:692-699.
- Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods*. 2018;15:233-234.
- Ramos LA, van der Steen WE, Sales Barros R, et al. Machine learning improves prediction of delayed cerebral ischemia in patients with subarachnoid hemorrhage. *J Neurointerv Surg*. 2019;11:497-502.
- Paliwal N, Jaiswal P, Tutino VM, et al. Outcome prediction of intracranial aneurysm treatment by flow diverters using machine learning. *Neurosurg Focus*. 2018;45:E7.
- Ueda D, Yamamoto A, Nishimori M, et al. Deep learning for MR angiography: automated detection of cerebral aneurysms. *Radiology*. 2019;290:187-194.

Conflict of interest statement: The authors declare that the article content was composed in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received 28 March 2019; accepted 29 June 2019

Citation: *World Neurosurg*. (2019) 131:e46-e51.

<https://doi.org/10.1016/j.wneu.2019.06.231>

Journal homepage: www.journals.elsevier.com/world-neurosurgery

Available online: www.sciencedirect.com

1878-8750/\$ - see front matter © 2019 Published by Elsevier Inc.