Elsevier Editorial System(tm) for Journal of Critical Care
Manuscript Draft

Manuscript Number: JCRC-D-13-00483

Title: Classification of Vasospasm in Aneurysmal Subarachnoid Hemorrhage Using Automated Data

Article Type: Research Paper

Keywords: Machine Learning, Artificial Intelligence, Subarachnoid Hemorrhage, Intracranial Vasospasm

Corresponding Author: Dr. Soojin Park, MD

Corresponding Author's Institution: Perelman School of Medicine, University of Pennsylvania

First Author: Alexander Roederer

Order of Authors: Alexander Roederer; John H Holmes, PhD; Michelle J Smith, MD; Insup Lee, PhD; Soojin Park, MD

Abstract: Purpose:
When vasospasm is detected after aneurysmal subarachnoid hemorrhage, it is treated with hypertensive or endovascular therapy. Current classification methods are resource-intensive, relying on specialty-trained professionals (nursing exams, transcranial dopplers, perfusion imaging). Passively obtained variables such as cerebrospinal fluid drainage volumes, sodium, glucose, blood pressure, and heart rate, have not been used to predict vasospasm. We hypothesize that these features may yield as much information as resource-intensive features to classify vasospasm.
Materials and Methods:
We studied 81 subarachnoid hemorrhage patients presenting within two days of onset. Vasospasm class (VSP) was defined by angiographic vasospasm warranting endovascular treatment. Naïve Bayes (NB) and logistic regression (LR) classifiers were trained on selected variable feature sets from the first three days of illness. Performance of trained classifiers was evaluated using area under the receiver operator characteristic curve (AUCclassifier) and F-measure (Fclassifier). Ablation analysis determined incremental utility of each variable and subsets.
Results:
43.2% developed VSP. NB classifier trained on all passively obtained features (AUCNB 0.708 and FNB 0.636) outperformed NB classifier trained on resource-intensive features (AUCNB 0.501 and FNB 0.349).
Conclusions:
Data-driven analysis of passively obtained clinical data predicted VSP better than current targeted resource-intensive monitoring techniques after aneurysmal subarachnoid hemorrhage. Automated classification of VSP may be possible.

Suggested Reviewers: Honglak Lee PhD
Assistant Professor, Computer Science & Engineering Division, University of Michigan, Ann Arbor
honglak@eecs.umich.edu

Dr Lee has research interests in Machine Learning and has done work in the area of stroke prediction. His knowledge of both the methodology and the clinical domain positions him well to review this paper.

Igor Kononenko PhD
Professor, Department of Artificial Intelligence, University of Ljubljana
xaigor@fri.uni-lj.si
Dr Kononenko area of research is machine learning in medicine, with an article titled: "Machine Learning for Medical Diagnosis: History, State of the Art and Perspective" that has been cited over 200 times.

Geert Meyfroidt MD
Associate Professor, Intensive Care Medicine, University Hospitals Leuven
geert.meyfroidt@uzleuven.be
Dr Meyfroidt's expertise in machine learning as well as intensive care medicine (specifically neurointensive care) positions him well to review this work.

To the Editor of Journal of Critical Care:

The manuscript has not been and will not be submitted elsewhere for publication. An abstract of preliminary findings was accepted by Neurocritical Care Society for poster presentation at their annual conference in Philadelphia, PA for October 2013.

The authors have no financial or personal relationships with people or organizations that could inappropriately influence this work.

Respectfully,

Soojin Park

Classification of Vasospasm in Aneurysmal Subarachnoid Hemorrhage Using Automated Data

Alexander Roederer (1), John H. Holmes PhD (2), Michelle J. Smith MD (3), Insup Lee PhD (1), Soojin Park MD (1,2,4,5)

1: Department of Computer and Information Science, School of Engineering and Applied Science, University of Pennsylvania

2: Department of Medical Informatics in Epidemiology, Center for Clinical Epidemiology & Biostatistics, Perelman School of Medicine, University of Pennsylvania

3: Department of Neurosurgery, Perelman School of Medicine, University of Pennsylvania

4: Department of Neurology, Perelman School of Medicine, University of Pennsylvania

5: Department of Anesthesiology & Critical Care, Perelman School of Medicine, University of Pennsylvania

Institution where work was done: Hospital of the University of Pennsylvania

Acknowledgements of Sources of Funding: This research was supported in part by the Penn Medicine Neuroscience Center (SP) and the National Science Foundation CNS-1035715 (AR, IL).

Address for correspondence: Soojin Park, 3 West Gates Building, 3400 Spruce Street, Philadelphia, PA 19104. Fax: (215) 614-1927, Phone: (215) 614-1533, Email: soojin.park@uphs.upenn.edu

Key words for indexing: Machine Learning, Artificial Intelligence, Subarachnoid Hemorrhage, Intracranial Vasospasm

Abstract

Purpose:

When vasospasm is detected after aneurysmal subarachnoid hemorrhage, it is treated with hypertensive or endovascular therapy. Current classification methods are resource-intensive, relying on specialty-trained professionals (nursing exams, transcranial dopplers, perfusion imaging). Passively obtained variables such as cerebrospinal fluid drainage volumes, sodium, glucose, blood pressure, and heart rate, have not been used to predict vasospasm. We hypothesize that these features may yield as much information as resource-intensive features to classify vasospasm.

Materials and Methods:

We studied 81 subarachnoid hemorrhage patients presenting within two days of onset. Vasospasm class (VSP) was defined by angiographic vasospasm warranting endovascular treatment. Naïve Bayes (NB) and logistic regression (LR) classifiers were trained on selected variable feature sets from the first three days of illness. Performance of trained classifiers was evaluated using area under the receiver operator characteristic curve ($AUC_{classifier}$) and F-measure ($F_{classifier}$). Ablation analysis determined incremental utility of each variable and subsets.

Results:

43.2% developed VSP. NB classifier trained on all passively obtained features (AUC_{NB} 0.708 and F_{NB} 0.636) outperformed NB classifier trained on resource-intensive features (AUC_{NB} 0.501 and F_{NB} 0.349).

Conclusions:

Data-driven analysis of passively obtained clinical data predicted VSP better than current targeted resource-intensive monitoring techniques after aneurysmal subarachnoid hemorrhage. Automated classification of VSP may be possible.

Introduction

Patients afflicted with aneurysmal subarachnoid hemorrhage (aSAH) are monitored in the intensive care unit for 10-14 days in order to monitor for neurologic and cardiac adverse sequelae. Delayed cerebral ischemia is the adverse event that most contributes to poor outcome, and has proven to have incomplete correlation with vasospasm (1). However, in the management of aSAH patients after the aneurysm is secured, clinicians are on alert for actionable events such as symptomatic angiographic vasospasm (2), which occurs most frequently 7 to 10 days after aneurysm rupture. When symptomatic vasospasm is detected in patients, they are treated with hypertensive therapy and in capable centers, localized intra-arterial delivery of vasodilators, angioplasty, or intra-aortic balloon pumps (2,3).

Frequent neurologic exams by trained nurses and healthcare professionals are vital for the detection of clinically significant vasospasm. Many centers will be alerted to preclinical vasospasm by daily transcranial doppler (TCD) and to delayed cerebral ischemia by perfusion scans (2). TCD is useful for detecting angiographic VSP in proximal segments of the intracranial arteries, especially the middle cerebral artery and basilar artery (4). TCD can be highly specific for angiographic vasospasm, but weakly sensitive (58.6%) (5). A few advanced centers additionally monitor for early (preclinical) ischemia using continuous electroencephalogram (cEEG) (6,7), beyond current guidelines. The trended information from these tests enable clinicians to continuously assess the likely classification of aSAH patients into categories of “in vasospasm” or “not

in vasospasm." The more days into the high-risk post-rupture period with or without suggestion of vasospasm, the more certain this classification becomes.

This method of vasospasm classification is heavily reliant on specialty-specific trained healthcare professionals. The sensitivity of neurologic exams is dependent on the expertise of the examiner. Furthermore, TCD are resource-intensive, dependent on the availability and workflow of skilled technicians; they are maximally available on a once daily basis, during normal workday hours.

Clinical features that are routinely monitored in aSAH patients but not traditionally used to detect symptomatic vasospasm include cerebrospinal fluid (CSF) drainage volumes, sodium values, intracranial pressures (ICP), blood pressures, and heart rates. We hypothesize that these passively obtained features hold as much information as traditional resource-intensive features (TCD or exam) to aid clinicians in classifying patients for symptomatic vasospasm.

Materials and Methods

We studied 89 high-grade aSAH patients admitted to the Neurocritical Care Unit at an academic medical center over an 11 year period (2001 to 2011). Approval for this study was obtained from the local Institutional Review Board.

Clinical Management of aSAH patients

Patients were managed according to a standard protocol (8, 9). An extraventricular drain was placed in all patients with Glasgow coma scale (GCS) <8, and in all patients with symptomatic hydrocephalus. Brain tissue oxygen (PbtO2) monitors

and parenchymal intracranial pressure (ICP) monitors were placed if the aSAH patient had a Fisher scale 3 or 4 (localized clots and/or vertical layers of blood 1 mm or greater in thickness or intracerebral or intraventricular clots) or GCS < 8. Goal PbtO2 was >20 and goal ICP was < 20. Arterial lines were placed in all patients, and systolic blood pressure (SBP) was maintained within 10% of baseline SBP if known, and otherwise between 120-160 with as little fluctuation as possible. Once the aneurysm was secured, SBP parameters were liberalized to 100-200 mm Hg (except for patients who underwent craniotomy, whose SBP were maintained < 160 mm Hg for 24 hours to minimize risk of postoperative bleeding). Hypertension was not treated unless SBP > 210 mm Hg. Nimodipine was administered enterally for 21 days. If SBP was >160 mm Hg, intravenous (IV) antihypertensive regimen was initiated (nicardipine or labetalol), while optimizing sedation and analgesia. Cerebral perfusion pressure (CPP) was maintained at > 60 mm Hg (this goal took precedence over SBP goal). Frank hypotension was avoided, using phenylephrine. A transthoracic echocardiogram was obtained if there was hypotension, suboptimal response to phenylephrine, ischemia or ST changes on ECG, or initial abnormal cardiac enzymes. If the ejection fraction was reduced or significant bradycardia existed, inotropic agents were used. A triple lumen venous catheter was placed in all patients, and euvolemia was maintained with a goal central venous pressure (CVP) of 6-12 mm Hg. Isotonic IV fluid was used, avoiding hypotonic IV fluid. Sodium was maintained within normal range of 133 to 143 mmol/L. Target blood sugar was 90-180 from 2001 to 2009, then 90-130 from 2009 to 2010, then 90-150 from 2010 to 2011. Seizure prophylaxis was given for 7 days after securing aneurysm. Statins, if present,

were not discontinued upon admission, and otherwise were started. Strict normothermia was maintained.

Vasospasm detection and management

Early detection and prevention of vasospasm and its consequences were a major focus of ICU care. Detection involved serial neurological exams and daily TCD studies including extracranial ICA velocity (for Lindegaard ratio). Patients were monitored every 1-3 hours for increased headache or clinical deterioration (focal neurologic change). Other diagnostic tests such as CT perfusion, Xenon CT, MR perfusion, and cEEG were available and used to help establish the presence of ischemia. Local practice for the optimal management of vasospasm included maintaining euvolemia, inducing hypertension and/or optimizing cardiac output. Available endovascular therapy included intra-arterial vasodilators (nicardipine) and balloon angioplasty.

Patient selection and Outcome definition

Inclusion criteria were aSAH patients with Fisher Grade 3 or 4 whose aneurysms were secured. Exclusion criteria were delayed presentation greater than 48 hours and early vasospasm (angiographically defined vasospasm that was treated with IA vasodilator or angioplasty during first 2 days of illness. The cohort of aSAH patients was classified into 2 groups (+VSP or -VSP) based on the occurrence of angiographic vasospasm treated with IA vasodilators (inclusive of all treatment with angioplasty). Patients without angiogram performed were included in the -VSP group.

TCD, Fisher grade, and nursing examination data were abstracted from the hospital records. Clinical data that are routinely monitored in ICU patients but not

traditionally used to detect symptomatic vasospasm were abstracted from the hospital record. These features included CSF drainage volumes, blood pressures, heart rates, ICP, glucose levels, and sodium levels.

Statistical methods:

Feature selection

Treatment (hypertensive therapy) often begins immediately before angiogram in suspected VSP cases. Because PBD 0-2 is a window of time before patient physiology may be altered through treatment, we focused on results utilizing only the first three days worth of data. Each clinical variable (CSF drainage volume, blood pressure, heart rate, ICP, glucose level, and sodium level) was tested for normality using both a Shapiro-Wilk test and visually, using a QQ plot. All features were normally distributed. Summary statistics (mean, median, minimum, maximum, standard deviation, and coefficient of variation) were calculated for each variable. Each statistic was calculated over each of the first three individual days of illness, as well as over the cumulative first three days of illness.

Using Student's t-tests, candidate features with a statistically significant difference ($p<0.05$) between the two VSP groups were selected for the final feature set. Classifiers were then constructed using this feature set.

Constructing the Classifiers

A naïve Bayes (NB) classifier and a logistic regression (LR) classifier were trained using the selected feature set. An NB classifier constructs a joint probability

distribution by "naively" assuming all features are independent, and applying Bayes' rule:

$$P(G|F = f) \propto P(G) \cdot P(f_1|G) \cdot P(f_2|G) \cdot \ldots \cdot P(f_n|G)$$

where G is the group (+VSP or -VSP) and f_i are the features.

This allowed us to learn a probability distribution over each of the groups, which could be used to calculate the probability a given new patient belongs to +VSP. An LR classifier, in contrast, attempts to fit a linear function over the feature values to the data by learning weights on each of the features, with a soft logistic function as the threshold for output:

$$P(F = f) = \frac{1}{1 + e^{-(w_0 + w_1 \cdot f_1 + w_2 \cdot f_2 + \cdots + w_n \cdot f_n)}}$$

where f_i are the features, and w_i are weights learned from the training data.

We used 10-fold cross validation to prevent over-fitting. In 10-fold cross validation, the data is randomly divided into 10 training subgroups of equal size. The first subgroup is selected and held in reserve while the remaining nine are used to train the classifier. The selected subgroup is then used to evaluate the classifier's performance. This process is repeated selecting each of the subgroups, and the results averaged across all runs. The Weka machine learning library (10) was used to construct and validate the classifiers.

NB and LR were chosen for use as they form a generative/discriminative classifier pair (that is, NB maximizes the total joint log-likelihood over the data, while

LR maximizes the total conditional likelihood over the same parametric model) (11), because they are easy to understand and are known to generalize relatively well. While patient vital signs are not expected to be independent, it has been shown that NB can perform optimally if dependencies distribute evenly in classes, thus cancelling each other out (12).

Evaluating Performance of the Classifiers

The performance of the trained classifiers was evaluated using two measures: area under the receiver operator characteristic (ROC) curve and the F-measure. The ROC is a plot that illustrates the performance of a binary classifier by varying its discrimination threshold and plotting the fraction of true positives over positive classifier results, versus the fraction of false positives over negative classifier results. The area under this curve, or c-statistic, is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one.

The F-measure is the harmonic mean of a classifier's recall and precision. Recall is the percentage of positive classifier results among the true positives:

$$recall = \frac{tp}{tp + tn}$$

It is equivalent to sensitivity. Precision is the percentage of true positives among the positive classifier results:

$$precision = \frac{tp}{tp + fp}$$

It is equivalent to positive predictive value. The F-measure is calculated as:

$$F = 2\left|\frac{precision \times recall}{precision + recall}\right|$$

Ablation Analysis

An ablation analysis was performed to determine the incremental value of each clinical variable in the model. Ablation analysis is a comparison of model performance when different subsets of features are included while training classifiers. Each of the relevant subsets of features was used individually to train both classifiers, and cross-validation was used to evaluate their performance with the metrics described above. For example, the classifiers were trained and tested using only features derived from TCD data. Then, the classifiers were trained and tested using only features derived from glucose values, and so on for each set of features described.

Results

89 patients with aSAH were examined; eight were excluded due to early vasospasm or low fisher grade. Of the remaining 81 patients, 35 (43.2%) were in the +VSP group. Patients were well-matched for clinical features except for age (mean age of +VSP group 56.4 (SD 14.4) vs –VSP group 50.8 (SD 9.3). Other clinical features examined were sex, tobacco use, clinical grade of SAH, and location of ruptured aneurysm. A table of patient characteristics is included in the electronic supplement (**Table E1**).

Post bleed day (PBD) 0 was defined as ictal day (day of rupture). Mean time of first angiographically defined VSP was PBD 6.02; range was PBD 2.43-12.72. 63

patients had 75 surveillance angiograms during PBD 0-2, to verify occlusion of aneurysm. None of these surveillance angiograms demonstrated +VSP.

In total, seven clinical variables were analyzed, including CSF output, BP, HR, ICP (measured parenchymally or extraventricularly), glucose value, and sodium value. A list showing average number of measurements per clinical variable per patient is included in the electronic supplement (**Table E2**). Summary statistics were performed; several features were deemed statistically significant ($p<0.05$). Results of the t-tests on the extracted summary statistics can be seen in **Table 1**.

Ablation analysis evaluated predictive performance of the summary statistics generated by each physiologic feature. These results can be found in **Table 2**. Several of the automated features, when used in isolation, produced classifiers which predicted VSP better than traditional measures. For example, a classifier trained on features derived from glucose values achieved an AUC_{NB} of 0.59 and an F_{NB} of 0.526. A classifier trained only on MAP achieved an AUC_{LR} of 0.526 and an F_{LR} of 0.645. More traditional measures, such as TCDs (AUC_{NB} 0.414, F_{NB} 0.29) and Fisher/Hunt Hess scores (AUC_{LR} 0.534, F_{LR} 0.235) did not perform as well (the combined models outperformed either of the two scores used in isolation). Exam scores (AUC_{NB} 0.595, F_{NB} 0.393) performed better.

The final ablation analysis considered larger subsets of features. An NB classifier trained on the automated feature set (CSF, MAP, HR, ICP, and glucose) achieved an AUC of 0.708 and an F-measure of 0.636, outperforming both an NB classifier built on

the traditional feature subset (AUC 0.501, F-measure 0.349) and an NB classifier built on all the features (combined traditional and automated) (AUC 0.625, F-measure 0.484).

Discussion

We have shown that an NB classifier using automated features of existing ICU data has the potential to classify VSP at least as well as targeted resource-heavy monitoring techniques after aSAH. It is important to note that the predictive performance of machine learning models built on physiologic feature summary statistics does not indicate anything about underlying causality or mechanism. Such models, however, may be useful as early warning or decision support systems, analyzing and integrating information in real-time to provide useful, targeted alerts to clinicians.

The advantages of an automated data-driven process in any domain are increased consistency and reduced human labor. In the healthcare domain, automation in the form of clinical decision support systems has the potential to allow early intervention and standardize quality of care while reducing clinicians' cognitive burden and preoccupation with algorithmic tasks, with the goal of improving outcome for patients. The methods we have used are not complex, and yet are not enabled by the current IT environment of most ICUs.

In our goal to identify opportunities for early decision support, we are limited in that this was a retrospective data set. Inherent in all retrospective data sets are possible confounding features that are not recorded. We attempted to mitigate the effect of these confounders by focusing our analysis around a specific cohort (high-grade aSAH patients). Also, our chosen "gold standard" (treated angiographic vasospasm) may be a

late marker for VSP, not taking into account early undocumented clinical suspicion. To explore the utility of our classifier further, we are now planning a study to compare its performance against prospectively recorded cognitive assessments and clinical decision making, while considering possible sources of confounding.

References (in sequential order)

(1) Vergouwen MD, Vermeulen M, van Gijn J, et al: Definition of delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage as an outcome event in clinical trials and observational studies: proposal of a multidisciplinary research group. Stroke 2010; 41:2391-2395.

(2) Connolly ES Jr, Rabinstein AA, Carhuapoma JR, et al. on behalf of the American Heart Association Stroke Council, Council on Cardiovascular Radiology and Intervention, Council on Cardiovascular Nursing, Council on Cardiovascular Surgery and Anesthesia, and Council on Clinical Cardiology: Guidelines for the management of aneurysmal subarachnoid hemorrhage: a guideline for healthcare professional from the American Heart Association/American Stroke Association. Stroke 2012; 43:1711-37.

(3) Spann RG, Lang DA, Birch AA, et al: Intra-aortic balloon counterpulsation: augmentation of cerebral blood flow after aneurysmal subarachnoid haemorrhage. Acta neurochirurgica 2001; 143:115-23.

(4) Sloan MA, Alexandrov AV, Tegeler CH, et al. on behalf of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology: Assessment: transcranial Doppler ultrasonography: report of the Therapeutics and

Technology Assessment Subcommittee of the American Academy of Neurology. Neurology 2004; 62:1468-81.

(5) Sloan MA, Haley Jr EC, Kassell NF, et al: Sensitivity and specificity of transcranial doppler ultrasonography in the diagnosis of vasospasm following subarachnoid hemorrhage. Neurology 1989; 39:1514-1518.

(6) Friedman D, Claassen J, Hirsch LJ: Continuous electroencephalogram monitoring in the intensive care unit. Anesth Analg 2009; 109:506-23.

(7) Claassen J, Taccone FS, Horn P, et al: Recommendations on the use of EEG monitoring in critically ill patients: consensus statement from the neurointensive care section of the ESICM. Intensive Care Med 2013; 39:1337-51.

(8) Mayberg MR, Batjer HH, Dacey R, et al: Guidelines for the management of aneurysmal subarachnoid hemorrhage: A statement for healthcare professionals from a special writing group of the Stroke Council, American Heart Association. Stroke 1994; 25:2315-28.

(9) Bederson JB, Connolly ES Jr, Batjer HH, et al: Guidelines for the management of aneurysmal subarachnoid hemorrhage: a statement for healthcare professionals from a special writing group of the Stroke Council, American Heart Association. Stroke 2009; 40:994-1025.

(10) Hall M, Frank E, Holmes G, et al: The WEKA data mining software: an update. SIGKDD Explorations 11(1):10-18, 2009 (newsletter).

(11) Ng AY, Jordan M: On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naïve Bayes. Neural Information Processing Systems 2001; 14:841-848.

(12) Zhang H: The Optimality of Naïve Bayes, in Barr V, Markov Z (eds): Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference FLAIRS. Miami Beach, FL, AAAI Press, 2004, pp 562-567.

Acknowledgements of Sources of Funding: This research was supported in part by the Penn Medicine Neuroscience Center (SP) and the National Science Foundation CNS-1035715 (AR, IL).

Conflicts of Interest: Authors have no personal or financial relationships with persons or organizations that would inappropriately influence this work.

Table 1

Table 1: T-test results. All summary statistics were calculated over one-day windows (PBD 0, PBD 1, PBD 2) and over a three-day window (PBD 0-2). Only statistically significant ($p < 0.05$) features are shown. Wilcoxon Rank-Sum tests are used instead of t-tests when comparing medians.

Feature	Window (PBD)	No VSP (Mean)	VSP (Mean)	p value	CI
Glucose Average	0-2	165.934	135.764	0.001	13.4-46.9
Glucose Average	1	166.823	141.532	0.012	5.8-44.7
Glucose Average	2	160.53	128.288	0.002	12.8-51.6
Glucose Max	0-2	220.41	185.231	0.03	3.5-66.9
Glucose Max	2	193.059	144.846	0.001	19.5-76.9
Glucose Median	0-2			0.001	11-46.5
Glucose Median	1			0.033	2-49
Glucose Median	2			0.013	5-43.5
Glucose Min	1	140.556	115.3	0.012	5.9-44.6
Glucose Min	2	134.353	114.538	0.035	1.5-38.2
HR CoeffVar	0	0.105	0.065	0.048	0.001-0.081
HR CoeffVar	0-2	0.141	0.091	0.006	0.016-0.085
HR CoeffVar	1	0.126	0.082	0.044	0.001-0.086
HR Min	0-2	58.394	69.571	0.038	-21.7-(-0.7)
HR StDev	0	9.161	4.761	0.024	0.67-8.13
HR StDev	0-2	11.355	7.272	0.006	1.29-6.88
ICP (Parenchymal) Max	1	34.632	18.4	0.044	0.5-32.0
ICP (Ventricle) Average	1	12.853	5.547	0.039	0.5-14.1
ICP (Ventricle) Median	1			0.044	0.00001-17
ICP (Ventricle) StDev	0-2	7.521	3.982	0.027	0.4-6.6
ICP (Ventricle) StDev	1	6.523	3.211	0.028	0.5-6.1
MAP Max	1	122.874	107.524	0.018	2.99-27.72
Sodium Average	1	142.972	139.4	0.034	0.32-6.82
Sodium CoeffVar	0-2	0.016	0.009	0.048	0.00007-0.013
Sodium Max	0-2	146.469	143.222	0.039	0.17-6.32
Sodium Max	1	144.321	140.2	0.03	0.46-7.78
Sodium StDev	0-2	2.347	1.335	0.037	0.065-1.959

Table 2

Table 2: Ablation Analysis. NB=naive Bayes classifier. LR=logistic regression classifier. AUC=area under the receiver operator curve. F=F-measure; weighted combination of accuracy and precision. All Automated Features=statistically significant summary features of CSF, MAP, HR, ICP, and Glucose. All Traditional Features=statistically significant summary features of TCDs, Exam Scores, and Fisher/Hunt Hess Grade.

Variable	AUC_{NB}	F_{NB}	AUC_{LR}	F_{LR}
TCDs	0.414	0.29	0.3	0.232
Exam Scores	0.595	0.393	0.336	0.405
Fisher & Hunt Hess Grades	0.499	0.163	0.534	0.235
CSF	0.475	0.259	0.538	0.353
MAP	0.647	0.182	0.526	0.645
HR	0.698	0.353	0.536	0.4
ICP	0.501	0.182	0.588	0.682
Glucose	0.59	0.526	0.535	0.424
Sodium	0.433	0.138	0.488	0.179
All Traditional Features	0.501	0.349	0.344	0.378
All Automated Features	0.708	0.636	0.544	0.513
All Features	0.625	0.484	0.459	0.475

Table E1

Table E1. Clinical Characteristics. P-value for age computed using student's t-test. All other p-values computed using Fisher's exact test. *Fisher-exact score not valid for sample sizes of 0. Statistics not applied to race because of high number of unknowns.

Characteristic	Class	+VSP	-VSP	p-value
Age (mean)		56.37 (SD 14.43)	50.83 (SD 9.33)	0.0420944
sex	female	35	22	0.2260508
	male	11	13	
race	asian	0	1	
	black	9	5	
	white	15	15	
	unknown	21	12	
Tobacco	Never smoker	10 (23.81%)	9 (28.12%)	0.9557217
	Current Smoker	18 (42.86%)	14 (43.75%)	
	Previous Smoker	4 (9.52%)	2 (6.25%)	
	Unknown	10 (23.81%)	7 (16.67%)	
Fisher	3	9 (20.45%)	9 (26.47%)	0.7189402
	4	25 (56.82%)	16 (47.06%)	
	3/4	10 (22.73%)	9 (26.47%)	
Hunt Hess	1	4 (9.3%)	3 (9.09%)	0.3257751
	2	4 (9.3%)	5 (15.15%)	
	3	8 (18.6%)	11 (33.33%)	
	4	19 (44.19%)	12 (36.36%)	
	5	8 (18.6%)	2 (6.06%)	
Ruptured Aneurysm	Anterior communicating artery	11 (25.00%)	15 (42.86%)	0.1640835
	Internal carotid artery	7 (15.91%)	4 (11.43%)	
	PCOM	10 (22.73%)	4 (11.43%)	
	Basilar artery	6 (13.64%)	3 (8.57%)	
	Middle Cerebral Artery	7 (15.91%)	6 (17.14%)	
	Other	3 (6.82%)	3 (8.57%)	

Table E2

Table E2: Average Number of Data Values by Clinical Variable. CSF output and ICP values only available 63 patients with extraventricular drains (64 had parenchymal ICP monitors, additionally).

Clinical Variable	Average # of Data Values per patient PBD 0-2 (SD)	Average # of Data Values per patient per day (SD)
CSF Output (24 Hrs Prior)	3.031 (2.061)	1.01 (0.687)
MAP	38.049 (40.697)	12.683 (13.566)
HR	38.58 (41.113)	12.86 (13.704)
ICP (extraventricular drain)	28.344 (33.024)	9.448 (11.008)
ICP (parenchymal)	22.219 (29.879)	7.406 (9.96)
Glucose Value	14.062 (17.763)	4.687 (5.921)
Sodium	1.852 (2.455)	0.617 (0.818)