

## Original Article

# Aneurysmal subarachnoid hemorrhage prognostic decision-making algorithm using classification and regression tree analysis

Benjamin W. Y. Lo, Hitoshi Fukuda<sup>1</sup>, Mark Angle, Jeanne Teitelbaum, R. Loch Macdonald<sup>2</sup>, Forough Farrokhyar<sup>3</sup>, Lehana Thabane<sup>4</sup>, Mitchell A. H. Levine<sup>5</sup>

Departments of Neurology and Neurosurgery, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada, <sup>1</sup>Departments of Neurosurgery, Kurashiki Central Hospital, Kyoto University, Okayama, Japan, <sup>2</sup>Division of Neurosurgery, St. Michael's Hospital, University of Toronto, Toronto, Ontario, Canada, <sup>3</sup>Department of Surgery, Department of Clinical Epidemiology and Biostatistics, <sup>4</sup>Department of Clinical Epidemiology and Biostatistics, <sup>5</sup>Department of Medicine, Division of Clinical Pharmacology, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada

E-mail: \*Benjamin W. Y. Lo - [benjamin.lo@mcgill.ca](mailto:benjamin.lo@mcgill.ca); Hitoshi Fukuda - [fukudaharpseal@gmail.com](mailto:fukudaharpseal@gmail.com); Mark Angle - [mark.angle@muhc.mcgill.ca](mailto:mark.angle@muhc.mcgill.ca); Jeanne Teitelbaum - [jteitelbaum@hotmail.com](mailto:jteitelbaum@hotmail.com); R. Loch Macdonald - [macdonaldlo@smh.ca](mailto:macdonaldlo@smh.ca); Forough Farrokhyar - [farrokh@mcmaster.ca](mailto:farrokh@mcmaster.ca); Lehana Thabane - [thabanl@mcmaster.ca](mailto:thabanl@mcmaster.ca); Mitchell A. H. Levine - [levinem@mcmaster.ca](mailto:levinem@mcmaster.ca)

\*Corresponding author

Received: 15 March 16 Accepted: 01 June 16 Published: 07 July 16

## Abstract

**Background:** Classification and regression tree analysis involves the creation of a decision tree by recursive partitioning of a dataset into more homogeneous subgroups. Thus far, there is scarce literature on using this technique to create clinical prediction tools for aneurysmal subarachnoid hemorrhage (SAH).

**Methods:** The classification and regression tree analysis technique was applied to the multicenter Tirilazad database (3551 patients) in order to create the decision-making algorithm. In order to elucidate prognostic subgroups in aneurysmal SAH, neurologic, systemic, and demographic factors were taken into account. The dependent variable used for analysis was the dichotomized Glasgow Outcome Score at 3 months.

**Results:** Classification and regression tree analysis revealed seven prognostic subgroups. Neurological grade, occurrence of post-admission stroke, occurrence of post-admission fever, and age represented the explanatory nodes of this decision tree. Split sample validation revealed classification accuracy of 79% for the training dataset and 77% for the testing dataset. In addition, the occurrence of fever at 1-week post-aneurysmal SAH is associated with increased odds of post-admission stroke (odds ratio: 1.83, 95% confidence interval: 1.56–2.45,  $P < 0.01$ ).

**Conclusions:** A clinically useful classification tree was generated, which serves as a prediction tool to guide bedside prognostication and clinical treatment decision making. This prognostic decision-making algorithm also shed light on the complex interactions between a number of risk factors in determining outcome after aneurysmal SAH.

**Key Words:** Aneurysmal subarachnoid hemorrhage, brain–body interactions, classification and regression tree analysis, prognostic decision making

### Access this article online

**Website:**

[www.surgicalneurologyint.com](http://www.surgicalneurologyint.com)

**DOI:**

10.4103/2152-7806.185786

**Quick Response Code:**

## INTRODUCTION

Neurologic outcome after aneurysmal subarachnoid hemorrhage (SAH) is influenced by a number of complex brain–body associations, including both non-treatment and treatment-related demographic, neurologic, and systemic factors. For bedside prognostication and clinical

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

**For reprints contact:** [reprints@medknow.com](mailto:reprints@medknow.com)

**How to cite this article:** Lo BW, Fukuda H, Angle M, Teitelbaum J, Macdonald RL, Farrokhyar F, et al. Aneurysmal subarachnoid hemorrhage prognostic decision-making algorithm using classification and regression tree analysis. *Surg Neurol Int* 2016;7:73.  
<http://surgicalneurologyint.com/Aneurysmal-subarachnoid-hemorrhage-prognostic-decision-making-algorithm-using-classification-and-regression-tree-analysis/>

decision making, a prediction tool is useful as an adjunct to complement clinical opinion and judgment.<sup>[8,9]</sup>

One of the most frequently used measures of global function outcomes after aneurysmal SAH is the Glasgow Outcome Score (GOS). It is graded into five subgroups, including good recovery, moderate disability but functional independence, functional dependence with severe disability, persistent vegetative state, and death. This scale can be easily and accurately administered, is reliable as self or surrogate reporting tool, and correlates well with neuropsychological test measures. Subtleties to changes in function can be captured by its extended version, as well as other neuropsychological and quality of life outcome scores.<sup>[6,16]</sup>

Classification and regression tree analysis is a statistical technique that involves creation of a decision tree by recursive partitioning of a dataset into more homogeneous subgroups. At each node, the best possible split for each variable is found by minimizing the impurity of subsequent daughter nodes. This technique prevents overfitting by pruning terminal nodes or leaves such that a decision tree is found with the smallest error rates. In addition, model performance can be enhanced by split sample validation techniques. Classification and regression tree analysis is a statistical technique that takes into account non-linear and high order interactions. It can handle large datasets with missing data, as well as high dimensionality.<sup>[1,2,11]</sup>

## Objectives

Using data from placebo and treatment arms of five clinical trials of Tirilazad, we aim to create a prognostic decision-making algorithm for aneurysmal SAH patients using the classification and regression tree analysis technique. In addition, model performance will be evaluated.

## MATERIALS AND METHODS

The Tirilazad database was used to investigate prognostic subgroups for outcome prediction in aneurysmal SAH patients. The Statistical Package for the Social Sciences (IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp) was used to create the classification and regression decision tree.

Candidate predictors included independent variables that reached a probability of 0.10 in univariate logistic regression analysis, including

1. Hospital admission neurological (Hunt and Hess) grade [1 – asymptomatic, 2 – headaches, cranial nerve palsy, 3 – mild focal deficit, 4 – stupor, moderate to severe hemiparesis, 5 – deep coma],
2. Age (years),
3. Time to surgery (hours),

4. Occurrence of post-admission fever on day 8,
5. Aneurysm size ( $\leq 12$  mm or  $> 12$  mm),
6. Presence of angiographic vasospasm on hospital admission,
7. Presence of intraventricular hemorrhage,
8. Development of cerebral edema,
9. Presence of intracerebral hemorrhage,
10. History of hypertension,
11. History of myocardial infarction,
12. Occurrence of stroke post-admission,
13. Prior episode of SAH,
14. Location of aneurysm,
15. Development of vasospasm during course of treatment, and
16. Development of seizures requiring antiepileptic medications.

The dependent variable used for the analysis was the dichotomized GOS at three months, where good outcome represents functional independence (GOS 5 and 4) and poor outcome represents functional dependence (GOS 3), persistent vegetative state (GOS 2), or death (GOS 1).

Statistical power was ensured by having effective sample size of more than 10 participants per variable, and at least 10 patients per final subgroup.<sup>[1,2,11,17]</sup> For preparing the decision tree, the Gini index splitting rule was used to decrease variance between classes through reduction of entropy where the largest categories were split into separate subgroups. Overfitting was avoided by using tree pruning.<sup>[1,2,11,17]</sup> Decision trees were generated with training and testing datasets, with percentages of favorable outcome reported in each terminal prognostic subgroup. Split sample analysis (70% training and 30% testing) was used to examine model performance, with reports of classification accuracy, risk estimate, and standard error for each decision tree. Univariate models using Chi-square analyses were performed to investigate the associations between explanatory variables generated in the classification trees.

## RESULTS

Patient selection and participation in the Tirilazad trials are summarized in Figure 1. In the Tirilazad patient populations, unfavorable outcome (functional dependence, persistent vegetative state, and death) at 3 months after aneurysmal SAH was observed in 1061 (30%) patients. Classification trees describing the relationships between neurological outcome and explanatory factors in aneurysmal SAH were generated using the training dataset [Figure 2] and testing dataset [Figure 3]. Figures 2 and 3 demonstrate seven terminal prognostic subgroups with the same node ranking order according to the percentages of unfavorable outcomes,

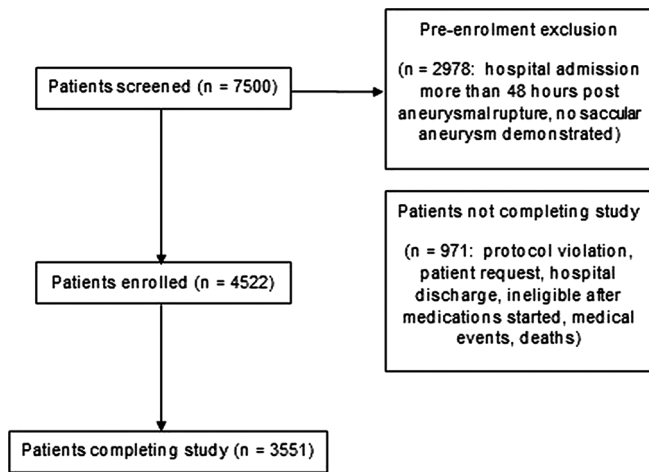


Figure 1: Patient selection and participation in the Tirilazad trials

with percentage differences between the two decision trees ranging from 1.3 to 7.8%. Seven terminal prognostic subgroups are illustrated as follows:

1. Good admission neurological grade (Hunt and Hess grades 1 or 2), absent post-admission stroke (Node #3)—8.7% unfavorable outcome (training dataset) and 7.4% unfavorable outcome (testing dataset),
2. Poor admission neurological grades (Hunt and Hess grades 3 or 4), absent post-admission stroke, absent post-admission fever (Node #10)—33.6% unfavorable outcome (training dataset) and 37.4% unfavorable outcome (testing dataset),
3. Good admission neurological grades, post-admission stroke, age <58.5 years (Node #7)—38.3% unfavorable outcome (training dataset) and 43.8% unfavorable outcome (testing dataset),

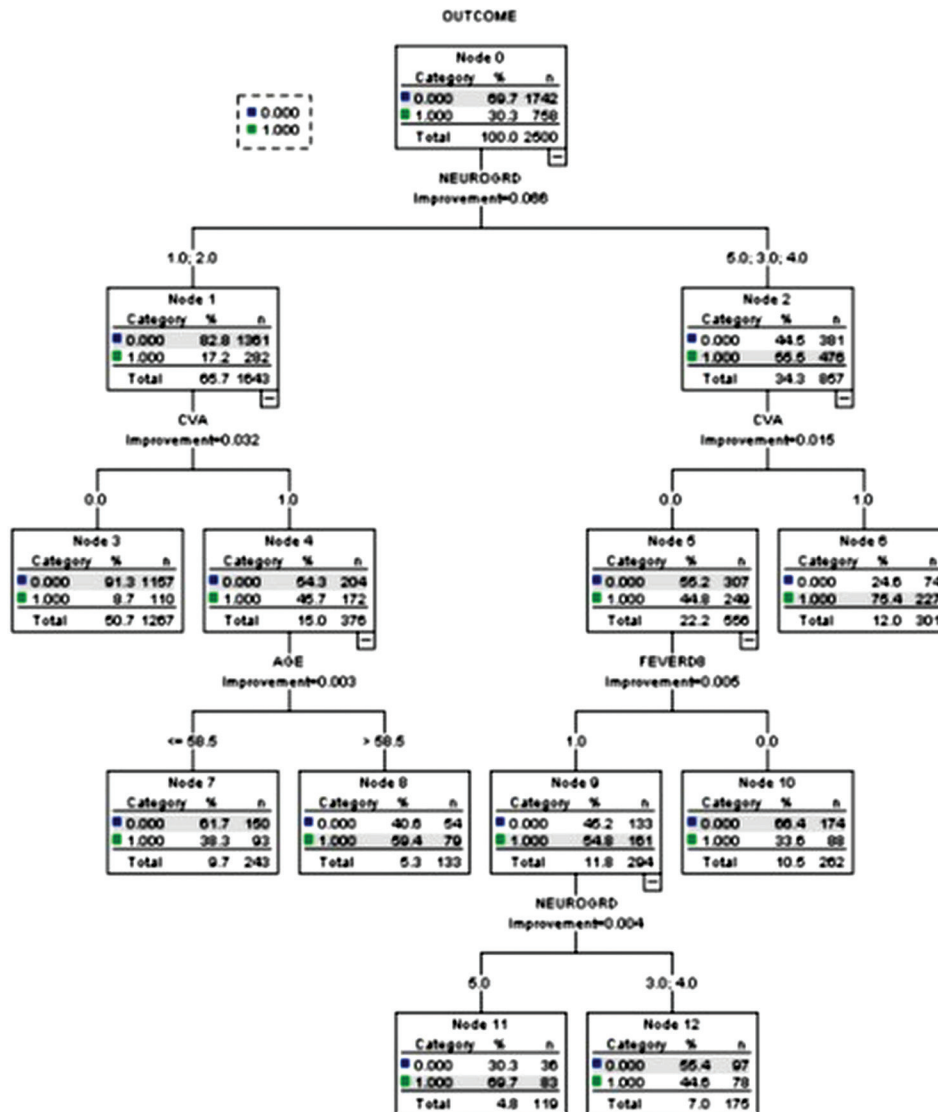
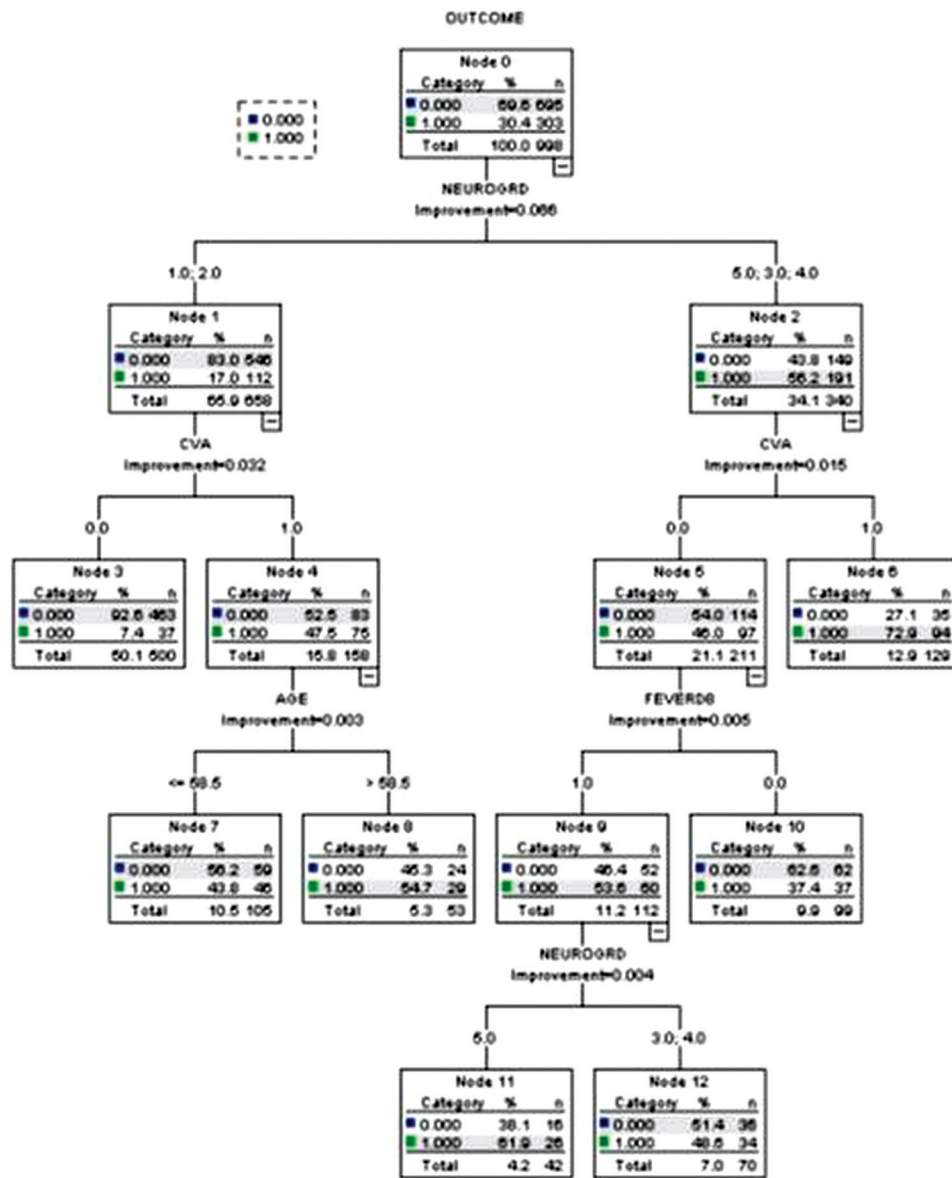


Figure 2: Aneurysmal subarachnoid hemorrhage classification tree using training data set. [For Node 0 (Outcome), 0 = favorable outcome [Glasgow Outcome Score (GOS) 5 and 4], 1 = unfavorable outcome (GOS 3, 2, and 1). For Nodes 1–12, 0 = absent, 1 = present.]



**Figure 3: Aneurysmal subarachnoid hemorrhage classification tree using testing data set. [For Node 0 (Outcome), 0 = favorable outcome [Glasgow Outcome Score (GOS) 5 and 4], 1 = unfavorable outcome (GOS 3, 2, and 1). For Nodes 1–12, 0 = absent, 1 = present.]**

4. Poor admission neurological grades, absent post-admission stroke, development of post-admission fever (Node #12)—44.6% unfavorable outcome (training dataset) and 48.6% unfavorable outcome (testing dataset),
5. Good admission neurological grades, post-admission stroke, age >58.5 years (Node #8)—59.4% unfavorable outcome (training dataset) and 54.7% unfavorable outcome (testing dataset),
6. Poor admission neurological grade (Hunt and Hess grade 5), absent post-admission stroke, development of post-admission fever (Node #11)—69.7% unfavorable outcome (training dataset) and 61.9% unfavorable outcome (testing dataset), and
7. Poor admission neurological grades, post-admission stroke (Node #6)—75.4% unfavorable outcome (training dataset) and 72.9% unfavorable outcome (testing dataset).

The classification and regression tree model generated using the training dataset had a classification accuracy of 79%, risk estimate of 0.21, with a standard error of 0.008. The model generated using the testing dataset had a classification accuracy of 77%, risk estimate of 0.23, and a standard error of 0.01. The occurrence of fever at 1 week post-aneurysmal rupture is associated with increased odds of post-admission stroke (odds ratio (OR): 1.83, 95% confidence interval (CI): 1.56–2.45,  $P < 0.01$ ).

Table 1 shows a clinical prognostic decision-making algorithm for aneurysmal SAH patients with prognostic subgroups based on the classification and regression tree derived from the Tirilzad database.

## DISCUSSION

Clinical prediction tools facilitate the process of prognostication and clinical decision making for both clinicians and patient families. The current classification and regression tree has seven terminal prognostic subgroups and makes use of the two most frequently retained clinical prognostic factors for long term neurologic outcome, namely, neurological grade<sup>[3,4,10,12,13,14,15]</sup> and age.<sup>[3,5,10,12,13,14,15]</sup> It also demonstrates the significance of both post-admission stroke and fever in outcome prediction.

In the present study, the occurrence of post-admission stroke increases the proportion of unfavorable neurologic outcome in aneurysmal SAH patients originally presenting with favorable admission neurological grades by 30%. In our prior analysis of the Tirilzad database, multivariable logistic regression analysis demonstrated that post-admission stroke increases the odds of poor neurological outcome in aneurysmal SAH patients by four fold (OR: 4.03, 95% CI: 2.11–7.69,  $P < 0.01$ ). Patients experiencing vasospasm are at an increased risk of post-admission strokes. In addition, several secondary injury cascade events may predispose these patients to post-admission strokes, including: (1) Microthrombi formation, (2) cortical spreading depression, (3) microvascular constriction, (4) proliferation of pro-inflammatory

cascade, (5) presence of blood–brain barrier disruption, and (6) inadequate collateral circulation.<sup>[7,9]</sup> Several contemporary neurocritical care strategies are used to decrease the likelihood of post-admission strokes. They include: (1) The use of milrinone, an inotropic vasodilator with anti-inflammatory properties, to prevent and treat vasospasm, (2) early decompressive craniectomy in patients with refractory increased intracranial pressures associated with cerebral edema, and (3) monitoring and treating seizures.

Fever is often a clinical indicator of neurological deterioration because it also triggers events in the secondary cascade of neurological injury. An epileptic aneurysmal SAH patient who develops post-admission fever has an increased odds of poor outcome by a factor of 2.4 (OR: 2.4, 95% CI: 1.86–3.06,  $P < 0.01$ ). The various causes of post-admission late onset fevers, including nosocomial infections, central neurological injury, thromboembolic events, and drug–drug interactions, can also lead to neurological complications, including increased intracranial pressures, cerebral edema, and post-admission strokes.<sup>[7,9]</sup> The occurrence of fever at 1 week post-hospital admission is associated with increased odds of post-admission strokes (OR: 1.83, 95% CI: 1.56–2.15,  $P < 0.01$ ), including vasospasm-induced delayed strokes. Aggressive symptomatic control and rigorous search for underlying etiology are, therefore, warranted.

## Limitations

The decision tree algorithm presented in this study was created using the Tirilzad database whereby prospective data was gathered to test an intervention, rather than for clinical prognostic purposes. Since the conduct

**Table 1: Clinical prognostic decision-making algorithm in aneurysmal subarachnoid hemorrhage, with prognostic subgroups based on classification and regression tree derived from the Tirilzad database**

Admission neurological grade	Age = 58.5 years	Post-admission fever	Post-admission stroke	Unfavorable outcome (%)	Unfavorable outcome category
Poor	Older	Present	Present	75.9	>75% very high
Poor	Older	Absent	Present	73.6	50-75% high
Poor	Younger	Present	Present	66.9	
Good	Older	Present	Present	63.2	
Poor	Older	Present	Absent	56.6	
Poor	Younger	Absent	Present	55.4	
Good	Older	Absent	Present	44.0	25-50% intermediate
Poor	Older	Absent	Absent	41.5	
Good	Younger	Present	Present	39.0	
Poor	Younger	Present	Absent	39.0	
Good	Younger	Absent	Present	25.7	
Good	Older	Present	Absent	18.5	<25% low
Poor	Younger	Absent	Absent	17.4	
Good	Older	Absent	Absent	14.3	
Good	Younger	Present	Absent	7.9	
Good	Younger	Absent	Absent	2.0	

of the clinical trials of Tirilazad, there have been advances in the surgical treatment and neurocritical care management of aneurysmal SAH patients. Despite variations in management at different centers, terminal prognostic subgroups depicted remain clinically relevant in a contemporary setting because they encompass frequently encountered factors including neurological grade, age, post-admission stroke, and fever. In this study, split validation technique was used to enhance model generalizability. Further investigations can make use of the classification and regression technique with external patient cohorts to further examine model generalizability.

## CONCLUSIONS

Clinical outcome after aneurysmal SAH may be influenced by interactions among a number of brain–body associations. Classification tree algorithm serves as a useful tool for prognostic decision making. The prognostic subgroups demonstrated the interplay of various underlying pathophysiologic mechanisms which, together, may adversely influence long-term neurologic outcome after aneurysmal SAH.

The classification tree generated in this article increases clinician and patient family awareness of the various demographic, neurologic, and systemic prognostic factors, which are pertinent for initial prognostication, namely, admission neurological grade and patient age. It also alerts the clinician regarding two important brain–body associations, including the occurrence of fever and delayed strokes, which may arise during hospitalization severely affecting patient outcomes. The clinician is, then, encouraged to tailor various treatment efforts to prevent and treat these, as well as other alterations in the brain–body interface in order to maximize the chances of survival and recovery after aneurysmal subarachnoid hemorrhage.

## Financial support and sponsorship

Nil.

## Conflicts of interest

There are no conflicts of interest.

## REFERENCES

- Breiman L. Classification and regression trees. Belmont, CA: Wadsworth; 1984.
- De'ath G, Fabricius K. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology* 2000;81:3178-92.
- Germanson T, Lanzino G, Kongable G, Torner J, Kassell N. Risk classification after aneurysmal subarachnoid hemorrhage. *Surg Neurol* 1998;49:155-63.
- Kahn J, Caldwell E, Deem S, Newell D, Heckbert S, Rubenfeld G. Acute lung injury in patients with subarachnoid hemorrhage: Incidence, risk factors and outcome. *Crit Care Med* 2006;34:196-202.
- Karamanakos PN, von Und Zu Fraunberg M, Bendel S, Huttunen T, Kurki M, Hernesniemi J, et al. Risk factors for three phases of 12-month mortality in 1657 patients from a defined population after acute aneurysmal subarachnoid hemorrhage. *World Neurosurg* 2012;78:631-9.
- Lo B, Stojic F, Spears J, Schweizer T, Macdonald R. Clinical outcomes, stroke trials and cognitive outcome. In: *The Behavioral Consequences of Stroke*. New York: Springer; 2014. pp. 283-314.
- Lo B, Fukuda H. Pathophysiology of brain somatic interactions in aneurysmal subarachnoid hemorrhage – Review and update. *Brain Disord Ther* 2015;4:183.
- Lo BW, Fukuda H, Nishimura Y, Farrokhyar F, Thabane L, Levine MA. Systematic review of clinical prediction tools and prognostic factors in aneurysmal subarachnoid hemorrhage. *Surg Neurol Int* 2015;6:135.
- Lo BW, Fukuda H, Nishimura Y, Macdonald RL, Farrokhyar F, Thabane L, et al. Pathophysiologic mechanisms of brain-body associations in ruptured brain aneurysms: A systematic review. *Surg Neurol Int* 2015;6:136.
- McGirt MJ, Woodworth GF, Ali M, Than KD, Tamargo RJ, Clatterbuck RE. Persistent perioperative hyperglycemia as an independent predictor of poor outcome after aneurysmal subarachnoid hemorrhage. *J Neurosurg* 2007;107:1080-5.
- Moisen G. Classification and regression trees. In: *Encyclopedia of Ecology*. Oxford, UK: Elsevier; 2008. pp. 582-8.
- Ogilvy CS, Cheung AC, Mitha AP, Hoh BL, Carter BS. Outcomes for surgical and endovascular management of intracranial aneurysms using a comprehensive grading system. *Neurosurgery* 2006;59:1037-42.
- Rabinstein AA, Friedman JA, Nichols DA, Pichelmann MA, McClelland RL, Manno EM, et al. Predictors of outcome after endovascular treatment of cerebral vasospasm. *AJNR Am J Neuroradiol* 2004;25:1778-82.
- Risselada R, Lingsma HF, Bauer-Mehren A, Friedrich CM, Molyneux AJ, Kerr RS, et al. Prediction of 60 day case-fatality after aneurysmal subarachnoid haemorrhage: Results from the International Subarachnoid Aneurysm Trial (ISAT). *Eur J Epidemiol* 2010;25:261-6.
- Rosengart AJ, Schultheiss KE, Tolentino J, Macdonald RL. Prognostic factors for outcome in patients with aneurysmal subarachnoid hemorrhage. *Stroke* 2007;38:2315-21.
- Sanchez C, Ogilvy C, Carter B. Outcomes studies in cerebrovascular neurosurgery. *Neurosurg Focus* 2007;22:E11.
- Yohannes Y, Webb P. Classification and regression trees, CART: A user manual for identifying indicators of vulnerability to famine and chronic food insecurity. Washington, D.C.: International Food Policy Research Institute, 1999.