Molecular Neurobiology (2019) 56:7128–7135
https://doi.org/10.1007/s12035-019-1601-7

Machine Learning Analysis of Matricellular Proteins and Clinical Variables for Early Prediction of Delayed Cerebral Ischemia After Aneurysmal Subarachnoid Hemorrhage

Satoru Tanioka[1] · Fujimaro Ishida[1] · Fumi Nakano[1] · Fumihiro Kawakita[2] · Hideki Kanamaru[2] · Yoshinari Nakatsuka[3] · Hirofumi Nishikawa[3] · Hidenori Suzuki[2] · pSEED group

Received: 9 February 2019 / Accepted: 3 April 2019 / Published online: 13 April 2019

Abstract

Although delayed cerebral ischemia (DCI) is a well-known complication after subarachnoid hemorrhage (SAH), there are no reliable biomarkers to predict DCI development. Matricellular proteins (MCPs) have been reported relevant to DCI and expected to become biomarkers. As machine learning (ML) enables the classification of various input data and the result prediction, the aim of this study was to construct early prediction models of DCI development with clinical variables and MCPs using ML analyses. Early-stage clinical data of 95 SAH patients in a prospective cohort were analyzed and applied to a ML algorithm, random forest, to construct three prediction models: (1) a model with only clinical variables on admission, (2) a model with only plasma levels of MCP (periostin, osteopontin, and galectin-3) at post-onset days 1–3, and (3) a model with both clinical variables on admission and MCP values at days 1–3. The prediction accuracy of the development of DCI, angiographic vasospasm, or cerebral infarction and the importance of each feature were computed. The prediction accuracy of DCI development was 93.9% in model 1, 87.2% in model 2, and 95.1% in model 3, but that of angiographic vasospasm or cerebral infarction was lower. The three most important features in model 3 for DCI were periostin, osteopontin, and galectin-3, followed by aneurysm location. All of the early-stage prediction models of DCI development constructed by ML worked with high accuracy and sensitivity. One-time early-stage measurement of plasma MCPs served for reliable prediction of DCI development, suggesting their potential utility as biomarkers.

Keywords Subarachnoid hemorrhage · Delayed cerebral ischemia · Matricellular protein · Machine learning · Prediction

Introduction

Delayed cerebral ischemia (DCI) is a well-known complication after aneurysmal subarachnoid hemorrhage (SAH) [1–6]. Cerebral vasospasm is not the only cause of DCI, and therefore, a multidisciplinary research group proposed the definition of DCI and angiographic vasospasm separately [1]. The precise pathogenesis of DCI remains unclear despite the prognostic value, and established biomarkers or reliable neuroimaging markers for DCI development are still unavailable except for the demonstration of angiographic vasospasm [1, 2].

Matricellular proteins (MCPs) are a component of extracellular matrix and play pivotal roles in various pathological conditions, such as inflammation, fibrosis, angiogenesis, vascular permeability, and cell death [7, 8]. Among many kinds of MCPs, osteopontin, periostin, and galectin-3 have been reported relevant to DCI, being expected to become valuable biomarkers and therapeutic targets of DCI [7, 9–12].

Medical researches using machine learning (ML) have been recently increasing [13–15]. In supervised ML algorithm, which is one of the ML algorithm categories, the operator provides the algorithm with pairs of inputs and desired outputs, and the algorithm discovers a way to produce the

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12035-019-1601-7) contains supplementary material, which is available to authorized users.

✉ Satoru Tanioka
satoru-tanioka@umin.net

[1] Department of Neurosurgery, Mie Chuo Medical Center, 2158-5 Myojin-cho, Hisai, Tsu, Mie 514-1101, Japan

[2] Department of Neurosurgery, Mie University Graduate School of Medicine, 2-174 Edobashi, Tsu, Mie 514-8507, Japan

[3] Department of Neurosurgery, Suzuka Kaisei Hospital, 112-1 Ko-cho, Suzuka, Mie 513-8505, Japan

desired output for a given input [16]: this method enables the classification of various input data and the prediction of results.

In this study, we applied clinical variables on admission and data of MCPs at post-onset days 1–3 as an input, and the development of DCI as the desired output to supervised ML algorithm, and aimed to construct early prediction models of DCI development. Additionally, the development of angiographic vasospasm and cerebral infarction was analyzed as the desired output in the same manner.

Methods

Study Population

Between 2013 and 2015, the prospective registry for searching mediators of neurovascular events after aneurysmal SAH (pSEED) was performed in seven tertiary hospitals in Japan (listed in Supplementary Appendix 1). The study was approved by the Institutional Ethics Committee, and written informed consent was obtained from the relatives. One hundred thirty-six patients were registered, and 95 patients met the following criteria: SAH on computed tomography (CT) scans or lumber puncture, modified Rankin Scale (mRS) 0–2 before onset, ≥ 20 years of age at onset, saccular aneurysm as the cause of SAH confirmed on CT angiography, digital subtraction angiography (DSA) or magnetic resonance (MR) angiography, aneurysm obliteration by coiling or clipping within 48 h of onset, and blood sampling on days 1–3 after onset (the day after coiling or clipping). Forty-one patients were excluded because of mRS ≥ 3 before onset ($n = 4$), ruptured dissecting aneurysm ($n = 6$), aneurysm obliteration at > 48 h of onset ($n = 8$), treatment by parent artery occlusion ($n = 3$), and incomplete data ($n = 20$) (Fig. 1).

Baseline clinical variables included age, sex, past medical history (SAH, cerebral infarction, hypertension, dyslipidemia, and diabetes mellitus), regular use of statin, smoking, alcohol, family history of SAH, pre-onset mRS, ruptured aneurysm location, admission World Federation of Neurological Surgeons (WFNS) grade, Fisher group, modified Fisher grade, and acute hydrocephalus.

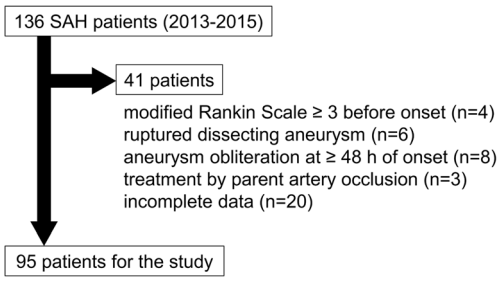

Fig. 1 A flow chart indicating the included and excluded patients in this study. SAH indicates subarachnoid hemorrhage

Inhospital Management

All patients were managed in the intensive care unit in accordance with the standards of SAH care in Japan: after aneurysm obliteration, control of hydrocephalus by cerebrospinal fluid drainage, intravenous injections of fasudil hydrochloride, nutritional management, and maintenance of euvolemia were performed. The selection of coiling or clipping and other medical management including administration of cilostazol, eicosapentaenoic acid, or statin for vasospasm prevention were decided by the onsite investigators.

Measurement of MCPs

After aneurysm obliteration, blood samples were collected with minimal stasis from a peripheral vein at post-onset days 1–3. All samples were centrifuged for 5 min at 3000*g*, and the supernatant fluid was stored at − 78 °C until assayed. The study group gathered the samples and blindly determined plasma concentrations of osteopontin, periostin, and galectin-3 using a commercially available enzyme-linked immunosorbent assay kit for osteopontin (code no. 27158; Immuno-Biological Laboratories, Fujioka, Japan), periostin (AG-45B-0004-KI01; AdipoGen Life Sciences, Liestal, Switzerland), and galectin-3 (code no. 27755; Immuno-Biological Laboratories, Fujioka, Japan).

Definition of DCI, Angiographic Vasospasm, and Cerebral Infarction

DCI was defined by the occurrence of new focal neurological impairments (such as hemiparesis, aphasia, apraxia, and/or neglect), and/or a decrease of at least 2 points on the Glasgow Coma Scale (either on the total score or on one of its individual components), which lasted for at least 1 h, and were not apparent immediately after aneurysmal obliteration. Potential causes of neurological impairments not assumed to be due to cerebral ischemia were rigorously excluded by means of clinical assessment, CT or MR imaging of the brain, and appropriate laboratory studies [1].

Angiographic vasospasm was defined as more than 50% reduction in the baseline vessel diameter of major cerebral arteries on CT angiography or DSA irrespective of symptoms [1, 9].

Cerebral infarction was defined by the development of cerebral infarction on CT or MR imaging of the brain within 6 weeks of SAH, which was neither presented on CT or MR imaging within 24–48 h after aneurysmal obliteration, nor attributable to causes other than DCI [1].

Machine Learning

Random forest, which is one of supervised ML algorithms, was used for the prediction analysis [16]. The data of clinical variables on admission and MCPs were applied to random forest, by which three prediction models were constructed: (1) a model with only clinical variables on admission, (2) a model with only MCPs, and (3) a model with both clinical variables on admission and MCPs. The development of DCI, angiographic vasospasm, or cerebral infarction was applied as the desired output. The synthetic minority over-sampling technique (SMOTE), in which the minority class is over-sampled by taking each minority class sample and creating synthetic examples along the line segments joining any or all of the k minority class nearest neighbors, was employed to redress the imbalance in the data distribution: depending on the amount of over-sampling required, neighbors from the k nearest neighbors were randomly chosen [17]. In the model construction process, the number of trees in random forest was determined according to the evaluation of DCI models using 50, 100, 150, and 200 to 1000 trees in increments of 100 trees. The Gini impurity was applied for the function to measure the quality of a split. The number of randomly selected features for the best split was set to the square root of the total number of the features in each model [18]. Leave-one-out (LOO) cross-validation was employed to evaluate generalization performance, in which a single data point was picked to be the test set for each split [16].

The accuracy was calculated by the number of correct predictions divided by the number of all samples [16]. The sensitivity was calculated by the number of true positive predictions divided by the number of all positive samples [16]. The importance of each feature was calculated with the mean decrease in the Gini impurity [19, 20]. In constructed models, the prediction accuracy of the development of DCI, angiographic vasospasm, or cerebral infarction and the importance of each feature were computed as the mean values of those for every split in LOO cross-validation.

The program Python (version 3.6.5, https://www.python.org/) and its libraries, NumPy (version 1.14.3, https://www.numpy.org/), scikit-learn (version 0.19.1, https://scikit-learn.org/stable/), imbalanced-learn (version 0.3.3, https://imbalanced-learn.org/en/stable/), and matplotlib (version 2.2.2, https://matplotlib.org/), were used for all data processing: the programming code was presented in Supplementary Appendix 2.

Statistical Analysis

Clinical variables on admission and MCPs were analyzed with SPSS statistics (IBM; Armonk, NY, USA). Student's t test, Fisher's exact test, and Pearson's chi-square test of independence were performed to assess statistical significance between two groups, appropriately. P values less than 0.05 were considered significant.

Results

Clinical variables on admission and MCP data at post-onset days 1–3 in models of DCI, angiographic vasospasm, and cerebral infarction are shown in Table 1 and Supplementary Tables S1 and S2, respectively. Inhospital variables in those models are shown in Table 2 and Supplementary Tables S3 and S4, respectively.

The number of trees in random forest was set to 100 in models 1 and 3, and 150 in model 2 for DCI; 200 in model 1, 100 in model 2, and 50 in model 3 for angiographic vasospasm; and 100 in model 1, and 200 in models 2 and 3 for cerebral infarction, with which the highest accuracy and sensitivity were achieved in each model (Fig. 2 and Supplementary Fig. S1 and S2). In LOO cross-validation, training was repeated 164 times in the DCI model, 128 times in the angiographic vasospasm model, and 142 times in the cerebral infarction model.

The mean values of the prediction accuracy of DCI development were 93.9% in model 1, 87.2% in model 2, and 95.1% in model 3. The sensitivity of DCI development was 92.6% in model 1, 94.7% in model 2, and 93.7% in model 3. The importance of each feature in DCI model 3 was shown in Fig. 3, in which the three most important features were periostin, osteopontin, and galectin-3, followed by aneurysm location.

The mean values of the prediction accuracy of angiographic vasospasm development were 73.4% in model 1, 72.7% in model 2, and 78.1% in model 3. The sensitivity of angiographic vasospasm development was 71.9% in model 1, 81.2% in model 2, and 76.6% in model 3. The importance of each feature in model 3 of angiographic vasospasm was shown in Supplementary Fig. S3, in which the three most important features were galectin-3, osteopontin, and age, followed by periostin.

The mean values of the prediction accuracy of cerebral infarction development were 81.7% in model 1, 78.9% in model 2, and 83.8% in model 3. The sensitivity of cerebral infarction development was 77.5% in model 1 and 84.5% in models 2 and 3. The importance of each feature in model 3 of cerebral infarction is shown in Supplementary Fig. S4, in which the three most important features were osteopontin, galectin-3, and periostin, followed by age.

Discussion

As ML algorithm, random forest was used in this study, which is one of the ensembles of decision trees: decision trees are widely used ML models for classification tasks, which learn a

Table 1 Clinical variables on admission and matricellular proteins at post-onset days 1–3 in the delayed cerebral ischemia (DCI) model

	DCI ($n = 13$)	No DCI ($n = 82$)	P value
Clinical variables			
Age	67.2 ± 11.9	65.2 ± 12.7	0.609[a]
Sex (male)	2 (15.4)	27 (32.9)	0.172[b]
Past history of SAH	1 (7.7)	4 (4.9)	0.529[b]
Past history of cerebral infarction	0 (0)	3 (3.7)	0.640[b]
Past history of hypertension	7 (53.8)	35 (42.7)	0.451[c]
Past history of dyslipidemia	2 (15.4)	10 (12.2)	0.515[b]
Past history of diabetes mellitus	1 (7.7)	3 (3.7)	0.451[b]
Regular use of statin	1 (7.7)	8 (9.8)	0.644[b]
Smoking	2 (15.4)	19 (23.2)	0.414[b]
Alcohol	3 (23.1)	26 (31.7)	0.393[b]
Family history of SAH	1 (7.7)	8 (9.8)	0.644[b]
Pre-onset modified Rankin Scale			0.453[b]
0	12 (92.3)	75 (91.5)	
1	0 (0)	5 (6.1)	
2	1 (7.7)	2 (2.4)	
Aneurysm location			0.562[b]
Anterior communicating artery	3 (23.1)	26 (31.7)	
Anterior cerebral artery	0 (0)	6 (7.3)	
Internal carotid artery	6 (46.2)	32 (39.0)	
Middle cerebral artery	4 (30.8)	10 (12.2)	
Posterior circulation	0 (0)	5 (6.1)	
Others	0 (0)	3 (3.7)	
Admission WFNS grade			0.968[b]
I	5 (38.5)	26 (31.7)	
II	3 (23.1)	23 (28.0)	
III	1 (7.7)	7 (8.5)	
IV	2 (15.4)	10 (12.2)	
V	2 (15.4)	16 (19.5)	
Fisher group			0.531[b]
2	0 (0)	7 (8.5)	
3	9 (69.2)	58 (70.7)	
4	4 (30.8)	17 (20.7)	
Modified Fisher grade			0.698[b]
1	0 (0)	7 (8.5)	
2	0 (0)	4 (4.9)	
3	7 (53.8)	45 (54.9)	
4	6 (46.2)	26 (31.7)	
Acute hydrocephalus	3 (23.1)	31 (37.8)	0.241[b]
Matricellular protein			
Osteopontin (pmol/l)	1040 ± 366	901 ± 350	0.192[a]
Periostin (pg/ml)	76,712 ± 23,363	61,026 ± 23,745	0.029[a]
Galectin-3 (pg/ml)	4252 ± 1403	4441 ± 1913	0.733[a]

Values are presented as n (%) or mean ± standard deviation. SAH, subarachnoid hemorrhage and WFNS, World Federation of Neurological Surgeons

[a] Student's t test between DCI and no DCI groups

[b] Fisher's exact test between DCI and no DCI groups

[c] Pearson's chi-square test between DCI and no DCI groups

Table 2 Inhospital variables in delayed cerebral ischemia (DCI) model

	DCI ($n = 13$)	No DCI ($n = 82$)	P value
Coiling	1 (7.7)	22 (26.8)	0.122[a]
Procedure-related complications			
Coiling			
Cerebral infarction	1 (7.7)	8 (9.8)	0.391[a]
Clipping			
Cerebral infarction	1 (7.7)	13 (15.9)	0.266[a]
Cerebral contusion	2 (15.4)	2 (2.4)	0.127[a]
Cerebrospinal fluid drainage			
Ventricular	2 (15.4)	22 (26.8)	0.307[a]
Cisternal	0 (0)	9 (11.0)	0.249[a]
Spinal	4 (30.8)	10 (12.2)	0.097[a]
Prevention of vasospasm			
Intravenous fasudil hydrochloride	13 (100)	80 (97.6)	0.744[a]
Cilostazol	11 (84.6)	69 (84.1)	0.665[a]
Eicosapentaenoic acid	6 (46.2)	34 (41.5)	0.750[b]
Statin	6 (46.2)	27 (32.9)	0.264[a]

Values are presented as n (%)

[a] Fisher's exact test between the DCI and no DCI groups

[b] Pearson's chi-square test between the DCI and no DCI groups

hierarchy of if/else questions, leading to a decision, and ensembles are methods which combine multiple ML models to construct more reliable models [16]. Because random forest does not require scaling of the input data and works well without heavy tuning of features, it is relatively easy-to-operate algorithm for clinicians. In addition, the importance of features can be calculated and ranked, facilitating the interpretation of constructed models. We consider that these are the advantages of random forest over other ML algorithms, such as support vector machines or artificial neural network [13–15].

As to MCPs measured in this study, the following has been shown in experimental studies: osteopontin was induced in damaged tissues and worked protectively against early brain injury (EBI) and DCI after SAH by preventing neuronal apoptosis and protecting blood-brain barrier [7, 21–25]; periostin was upregulated in brain after SAH and responsible for EBI through activating downstream signaling pathways and interacting with another MCP tenascin-C [7, 26]; and galectin-3 was secreted by brain capillary endothelial cells, acted as a Toll-like receptor 4 ligand and caused EBI [7, 10, 27]. In addition, individual predictive impacts of the MCPs on DCI development or eventual poor outcomes have been reported in prospective clinical studies [9, 10, 28]. EBI is an acute pathophysiological event that occurs within the first 72 h of SAH, which is induced by elevated intracranial pressure and subsequent transient global cerebral ischemia as well as subarachnoid blood degradation products due to aneurysmal rupture, and is considered to have a central role in the development of DCI [7].

In regard to factors affecting DCI development, age, aneurysm location, admission WFNS grade, and modified Fisher Scale have been reported in large-scale studies [2–6].

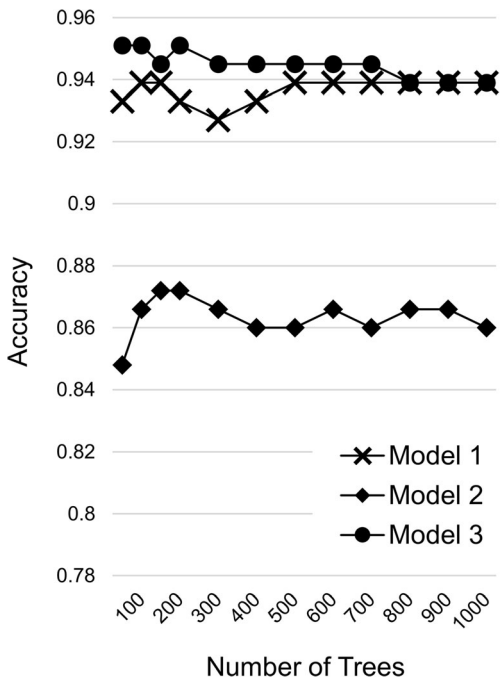

Fig. 2 Accuracy of each DCI model with 50, 100, 150, and 200 to 1000 trees in increments of 100 trees

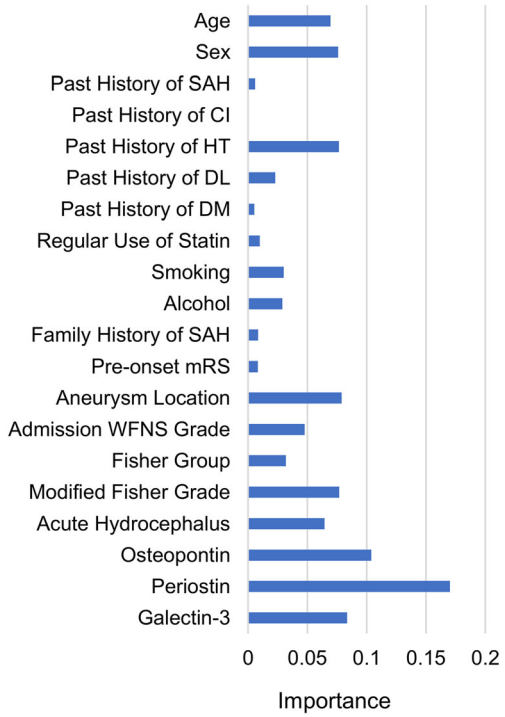

Fig. 3 The importance of each feature in a delayed cerebral ischemia model with combining matricellular proteins and clinical variables on admission. CI, cerebral infarction; DL, dyslipidemia; DM, diabetes mellitus; HT, hypertension; mRS, modified Rankin Scale; SAH, subarachnoid hemorrhage; and WFNS, World Federation of Neurological Surgeons

However, it is controversial whether younger or older age is associated with a higher incidence of DCI [2–6, 29, 30], although age was an important factor for DCI development in this study. Since the definition of DCI is inconsistent among studies, the interpretation of the findings is by no means easy [2–6]. In this study, the definition proposed by Vergouwen et al. [1] was adopted, in which the assessment of DCI was based on the occurrence of clinical deterioration, whereas that of cerebral infarction was based on neuroimaging, such as CT or MR imaging. In contrast, in some of previously reported large-scale studies, cerebral infarction and vasospasm were included in the definition of DCI [2–6, 31]. Hence, a simple comparison between the present and previous studies cannot be made.

All models for the prediction of DCI development worked with high accuracy and sensitivity in the ML analyses in this study. Since this study used a prospectively maintained database, we consider that these models have sufficient quality to be practically used if the number of patients is increased. Although the prediction accuracy in model 3 was the highest, it is notable that models 1 and 2 also had high accuracy. The high accuracy of model 1 may indicate that clinical data on admission serve for reliable prediction as reported in previous studies [2–6]. Meanwhile, the high accuracy of model 2 that was constructed with only MCP data at post-onset days 1–3 may indicate that the combination of MCP measurements at an acute stage becomes useful biomarkers for DCI development, possibly by reflecting the extent of EBI. Although individual MCPs may have predictive values for DCI development as described above, it may be difficult that a single MCP can be a clinically reliable and effective biomarker for DCI development. This is because plasma concentrations of each MCPs change with time after SAH and may reflect a variety of pathological processes underlying DCI, angiographic vasospasm, and cerebral infarction as shown in this study [7, 9, 10]. However, random forest, which constructs models by combining various features and excels in evaluation of combined features, showed the possibility that measurements of MCPs at an acute stage of SAH are enough to predict DCI development in this study. Moreover, it is noteworthy that the sensitivity was the highest in model 2 using MCP data only, enhancing the potential utility of MCPs. This is the first study to construct prediction models of DCI development by ML and to show that plasma values of periostin, osteopontin, and galectin-3 at an acute stage of SAH may be the most important to predict DCI among diverse clinical variables including well-known predictors of WFNS and Fisher grades. Measurements of plasma levels of the MCPs at post-onset days 1–3 may be useful for early identification of patients at risk for DCI to apply an aggressive DCI treatment strategy. If the ML platform is advanced and easily available for use, our constructed models can be widely available to clinicians.

In comparison with DCI models, models for angiographic vasospasm and cerebral infarction showed lower accuracy. The findings may indicate that other clinical variables at an acute stage or later stages are more important for the development of cerebral vasospasm and infarction. It is known that angiographic vasospasm does not always cause DCI and cerebral infarction, but does in association with various factors including genetic variations, physiological derangement, and microcirculatory disturbance: on the contrary, DCI can develop without vasospasm [32, 33].

In the DCI model 3, the three most important features were MCPs. However, the accuracy was higher in model 1, which contained no MCPs, compared with model 2. The precise interpretation is difficult, but the following explanation is possible: the number of features for model construction in model 2 was small, and therefore, the variation of decision trees was restricted, causing less accuracy. Anyway, the assessment based on random forest helps understanding of the importance of each feature in models. However, it should be acknowledged that it does not mean the meaningless of a feature even if the feature has a low value in importance: it only means that the feature was not picked by the decision tree, likely because another feature encoded the same information [16]. However, since we performed thorough cross-validation in this study, the feature importance having high relevance may be correct.

There are some limitations in this study. First, more study patients are needed to allow models to have sufficient quality for practical use. Half of excluded patients lacked some of data, most of which were blood sampling data: this might affect the accurate model construction. It is difficult to determine an appropriate sample size in ML analyses, which differs according to the quality of input data. However, we should make an effort to increase sample size and confirm that prediction accuracy and its stability have reached a plateau irrespective of an increase in sample size [34]. Second, only MCPs were investigated as biomarkers in this study. There are many biomarkers including microRNAs reported to have the relationships with DCI, angiographic vasospasm, and cerebral infarction [35, 36]. Thus, future studies should include other potentially important biomarkers. Third, over-sampling was conducted because of the imbalanced number of patients in desired output in all models. Although the SMOTE has been employed in several medical studies [37, 38], we should recognize that synthetic samples are generated in this method, which may affect the results. Last, although random forest was adopted for ML algorithm in this study, artificial neural network or other algorithms might have developed precise prediction [1, 2]. However, we consider that sophisticated data collection is much more important than ML algorithm selection to construct the practical model.

Conclusions

We constructed three early prediction models of DCI development by ML combining the data of MCPs at post-onset days 1–3 and clinical variables on admission in a prospective cohort and revealed that all models worked with high accuracy. It is notable that one-time early-stage measurement of MCPs served for reliable prediction of DCI development, indicating the potential utility of MCPs as biomarkers.

Acknowledgements We thank Dr. Nobuhisa Kashiwagi (the Institute of Statistical Mathematics, Tokyo, Japan) for review of statistical analyses.

Funding This work was supported by a grant-in-aid for Scientific Research from Japan Society for the Promotion of Science to Dr. HS (grant number 17K10825).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

1. Vergouwen MDI, Vermeulen M, van Gijn J, Rinkel GJE, Wijdicks EF, Muizelaar JP, Mendelow AD, Juvela S et al (2010) 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 41:2391–2395. https://doi.org/10.1161/STROKEAHA.110.589275
2. Foreman PM, Chua MH, Harrigan MR, Fisher WS, Tubbs RS, Shoja MM, Griessenauer CJ (2017) External validation of the practical risk chart for the prediction of delayed cerebral ischemia following aneurysmal subarachnoid hemorrhage. J Neurosurg 126:1530–1536. https://doi.org/10.3171/2016.1.JNS152554
3. de Rooij NK, Greving JP, Rinkel GJE, Frijns CJM (2013) Early prediction of delayed cerebral ischemia after subarachnoid hemorrhage: development and validation of a practical risk chart. Stroke 44:1288–1294. https://doi.org/10.1161/STROKEAHA.113.001125
4. Lee H, Perry JJ, English SW, Alkherayf F, Joseph J, Nobile S, Zhou LL, Lesiuk H et al (2018) Clinical prediction of delayed cerebral ischemia in aneurysmal subarachnoid hemorrhage. J Neurosurg. https://doi.org/10.3171/2018.1.JNS172715 [published Online First: 8 June 2018]
5. de Oliveira Manoel AL, Jaja BN, Germans MR, Yan H, Qian W, Kouzmina E, Marotta TR, Turkel-Parrella D et al (2015) The VASOGRADE: a simple grading scale for prediction of delayed cerebral ischemia after subarachnoid hemorrhage. Stroke 46:1826–1831. https://doi.org/10.1161/STROKEAHA.115.008728
6. Crobeddu E, Mittal MK, Dupont S, Wijdicks EFM, Lanzino G, Rabinstein AA (2012) Predicting the lack of development of delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage. Stroke 43:697–701. https://doi.org/10.1161/STROKEAHA.111.638403
7. Suzuki H, Nishikawa H, Kawakita F (2018) Matricellular proteins as possible biomarkers for early brain injury after aneurysmal subarachnoid hemorrhage. Neural Regen Res 13:1175–1178. https://doi.org/10.4103/1673-5374.235022
8. Murphy-Ullrich JE, Sage EH (2014) Revisiting the matricellular concept. Matrix Biol 37:1–14. https://doi.org/10.1016/j.matbio.2014.07.005
9. Nakatsuka Y, Shiba M, Nishikawa H et al (2018) Acute-phase plasma osteopontin as an independent predictor for poor outcome after aneurysmal subarachnoid hemorrhage. Mol Neurobiol 55:6841–6849. https://doi.org/10.1007/s12035-018-0893-3
10. Nishikawa H, Nakatsuka Y, Shiba M et al (2018) Increased plasma galectin-3 preceding the development of delayed cerebral infarction and eventual poor outcome in non-severe aneurysmal subarachnoid hemorrhage. Transl Stroke Res 9:110–119. https://doi.org/10.1007/s12975-017-0564-0
11. Nishikawa H, Suzuki H (2017) Implications of periostin in the development of subarachnoid hemorrhage-induced brain injuries. Neural Regen Res 12:1982–1984. https://doi.org/10.4103/1673-5374.221150
12. Nishikawa H, Suzuki H (2018) Possible role of inflammation and galectin-3 in brain injury after subarachnoid hemorrhage. Brain Sci 8. https://doi.org/10.3390/brainsci8020030
13. Brusko GD, Kolcun JPG, Wang MY (2018) Machine-learning models: the future of predictive analytics in neurosurgery. Neurosurgery 83:E3–E4. https://doi.org/10.1093/neuros/nyy166
14. Muhlestein WE, Akagi DS, Davies JM, Chambless LB (2018) Predicting inpatient length of stay after brain tumor surgery:

developing machine learning ensembles to improve predictive performance. Neurosurgery. https://doi.org/10.1093/neuros/nyy343 [published Online First: 3 August 2018]
15. Senders JT, Arnaout O, Karhade AV, Dasenbrock HH, Gormley WB, Broekman ML, Smith TR (2018) Natural and artificial intelligence in neurosurgery: a systematic review. Neurosurgery 83: 181–192. https://doi.org/10.1093/neuros/nyx384
16. Müller AC, Guido S (2017) Introduction to machine learning with Python: a guide for data scientists. 1st ed. O'Reilly, Sebastopol
17. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357. https://doi.org/10.1613/jair.953
18. Random forest classifier. scikit-learn. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html. Accessed 8 Feb 2019
19. Louppe G (2015) Understanding random forests: from theory to practice. arXiv. https://arxiv.org/abs/1407.7502. Accessed 8 Feb 2019
20. Ensemble methods. scikit-learn. https://scikit-learn.org/stable/modules/ensemble.html. Accessed 8 Feb 2019
21. Suzuki H, Hasegawa Y, Kanamaru K, Zhang JH (2010) Mechanisms of osteopontin-induced stabilization of blood-brain barrier disruption after subarachnoid hemorrhage in rats. Stroke 41:1783–1790. https://doi.org/10.1161/STROKEAHA.110.586537
22. Suzuki H (2015) What is early brain injury? Transl Stroke Res 6:1–3. https://doi.org/10.1007/s12975-014-0380-8
23. Suzuki H, Ayer R, Sugawara T, Chen W, Sozen T, Hasegawa Y, Kanamaru K, Zhang JH (2010) Protective effects of recombinant osteopontin on early brain injury after subarachnoid hemorrhage in rats. Crit Care Med 38:612–618. https://doi.org/10.1097/CCM.0b013e3181c027ae
24. Suzuki H, Hasegawa Y, Chen W, Kanamaru K, Zhang JH (2010) Recombinant osteopontin in cerebral vasospasm after subarachnoid hemorrhage. Ann Neurol 68:650–660. https://doi.org/10.1002/ana.22102
25. Topkoru BC, Altay O, Duris K, Krafft PR, Yan J, Zhang JH (2013) Nasal administration of recombinant osteopontin attenuates early brain injury after subarachnoid hemorrhage. Stroke 44:3189–3194. https://doi.org/10.1161/STROKEAHA.113.001574
26. Liu L, Kawakita F, Fujimoto M, Nakano F, Imanaka-Yoshida K, Yoshida T, Suzuki H (2017) Role of periostin in early brain injury after subarachnoid hemorrhage in mice. Stroke 48:1108–1111. https://doi.org/10.1161/STROKEAHA.117.016629
27. Nishikawa H, Liu L, Nakano F, Kawakita F, Kanamaru H, Nakatsuka Y, Okada T, Suzuki H (2018) Modified citrus pectin prevents blood-brain barrier disruption in mouse subarachnoid hemorrhage by inhibiting galectin-3. Stroke 49:2743–2751. https://doi.org/10.1161/STROKEAHA.118.021757
28. Kanamaru H, Kawakita F, Nakano F et al (2019) Plasma periostin and delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage. Neurotherapeutics. https://doi.org/10.1007/s13311-018-00707-y
29. Yoshimoto Y, Kwak S (1995) Age-related multifactorial causes of neurological deterioration after early surgery for aneurysmal subarachnoid hemorrhage. J Neurosurg 83:984–988. https://doi.org/10.3171/jns.1995.83.6.0984
30. Lanzino G, Kassell NF, Germanson TP, Kongable GL, Truskowski LL, Torner JC, Jane JA (1996) Age and outcome after aneurysmal subarachnoid hemorrhage: why do older patients fare worse? J Neurosurg 85:410–418. https://doi.org/10.3171/jns.1996.85.3.0410
31. Macdonald RL, Kassell NF, Mayer S, Ruefenacht D, Schmiedek P, Weidauer S, Frey A, Roux S et al (2008) Clazosentan to overcome neurological ischemia and infarction occurring after subarachnoid hemorrhage (CONSCIOUS-1): randomized, double-blind, placebo-controlled phase 2 dose-finding trial. Stroke 39:3015–3021. https://doi.org/10.1161/STROKEAHA.108.519942
32. Rosalind Lai PM, Du R (2015) Role of genetic polymorphisms in predicting delayed cerebral ischemia and radiographic vasospasm after aneurysmal subarachnoid hemorrhage: a meta-analysis. World Neurosurg 84:933–941.e2. https://doi.org/10.1016/j.wneu.2015.05.070
33. Suzuki H, Shiba M, Nakatsuka Y, Nakano F, Nishikawa H (2017) Higher cerebrospinal fluid pH may contribute to the development of delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage. Transl Stroke Res 8:165–173. https://doi.org/10.1007/s12975-016-0500-8
34. Cui Z, Gong G (2018) The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. Neuroimage 178:622–637. https://doi.org/10.1016/j.neuroimage.2018.06.001
35. Chan MTH, Wong JYY, Leung AKT et al (2019) Plasma and CSF miRNA dysregulations in subarachnoid hemorrhage reveal clinical courses and underlying pathways. J Clin Neurosci 62:155–161. https://doi.org/10.1016/j.jocn.2018.11.038
36. Jabbarli R, Pierscianek D, Darkwah Oppong M, Sato T, Dammann P, Wrede KH, Kaier K, Köhrmann M et al (2018) Laboratory biomarkers of delayed cerebral ischemia after subarachnoid hemorrhage: a systematic review. Neurosurg Rev. https://doi.org/10.1007/s10143-018-1037-y [published Online First: 10 October 2018]
37. Alghamdi M, Al-Mallah M, Keteyian S et al (2017) Predicting diabetes mellitus using SMOTE and ensemble machine learning approach: The Henry Ford ExercIse Testing (FIT) roject. PLoS One 12:e0179805. https://doi.org/10.1371/journal.pone.0179805
38. Sui Y, Wei Y, Zhao D (2015) Computer-aided lung nodule recognition by SVM classifier based on combination of random undersampling and SMOTE. Comput Math Methods Med 368674:1–13. https://doi.org/10.1155/2015/368674