



OPEN

Development of postoperative delirium prediction models in patients undergoing cardiovascular surgery using machine learning algorithms

Chie Nagata^{1✉}, Masahiro Hata², Yuki Miyazaki², Hirotada Masuda³, Tamiki Wada², Tasuku Kimura⁴, Makoto Fujii¹, Yasushi Sakurai⁴, Yasuko Matsubara⁴, Kiyoshi Yoshida¹, Shigeru Miyagawa³, Manabu Ikeda² & Takayoshi Ueno^{1,3}

Associations between delirium and postoperative adverse events in cardiovascular surgery have been reported and the preoperative identification of high-risk patients of delirium is needed to implement focused interventions. We aimed to develop and validate machine learning models to predict post-cardiovascular surgery delirium. Patients aged ≥ 40 years who underwent cardiovascular surgery at a single hospital were prospectively enrolled. Preoperative and intraoperative factors were assessed. Each patient was evaluated for postoperative delirium 7 days after surgery. We developed machine learning models using the Bernoulli naive Bayes, Support vector machine, Random forest, Extra-trees, and XGBoost algorithms. Stratified fivefold cross-validation was performed for each developed model. Of the 87 patients, 24 (27.6%) developed postoperative delirium. Age, use of psychotropic drugs, cognitive function (Mini-Cog < 4), index of activities of daily living (Barthel Index < 100), history of stroke or cerebral hemorrhage, and eGFR (estimated glomerular filtration rate) < 60 were selected to develop delirium prediction models. The Extra-trees model had the best area under the receiver operating characteristic curve (0.76 [standard deviation 0.11]; sensitivity: 0.63; specificity: 0.78). XGBoost showed the highest sensitivity (AUROC, 0.75 [0.07]; sensitivity: 0.67; specificity: 0.79). Machine learning algorithms could predict post-cardiovascular delirium using preoperative data.

Trial registration: UMIN-CTR (ID; UMIN000049390).

Delirium is a variable acute-onset disturbance of attention and consciousness that has multiple etiologies¹. Surgery is a common risk factor for delirium, with postoperative delirium occurring in 14–24% of patients following non-cardiac surgery^{2,3}. Furthermore, approximately 26–56% of cardiovascular surgery patients who require the use of cardiopulmonary bypass (CPB) develop postoperative delirium^{4–7}. Delirium following cardiovascular surgery is independently associated with a decline in cognitive function⁸, activities of daily living (ADL)^{6,9,10} and increased mortality¹¹. Thus, postoperative delirium is an important factor that influences patient prognosis.

The American Heart Association issued a statement in 2020 to integrate measures against delirium into Cardiovascular Intensive Care Unit (ICU) care¹². Recently, the European Society of Anesthesiology (2017)¹³ and POQI-6 (2019)¹⁴ have introduced guidelines for postoperative prevention and recommended multicomponent non-pharmacological interventions for delirium. The Hospital Elder Life Program (HELP)¹⁵, the most well-known non-pharmacological multidisciplinary program, reduces delirium by 53%¹⁶. However, it is difficult to apply these measures to all patients¹⁴. In a cross-sectional study of acute hospitals implementing HELP, the most common reasons for non-compliance with the original HELP program were lack of adequate staffing

¹Division of Health Sciences, Osaka University Graduate School of Medicine, 1-7 Yamadaoka, Suita, Osaka 565-0871, Japan. ²Department of Psychiatry, Osaka University Graduate School of Medicine, Suita, Osaka, Japan. ³Department of Cardiovascular Surgery, Osaka University Graduate School of Medicine, Suita, Osaka, Japan. ⁴SANKEN (The Institution of Scientific and Industrial Research), Osaka University, Ibaraki, Osaka, Japan. ✉email: nagatachie1994@sahs.med.osaka-u.ac.jp

and budgetary limitation¹⁷. As pharmacological therapy, ramlteon and suvorexant also reportedly reduce delirium^{18,19}, but premedication to prevent delirium is not recommended for use in all patients²⁰. Therefore, early prediction and identification of individuals at high risk for delirium are needed to provide targeted and efficient interventions^{14,21}.

Several studies have aimed to predict postoperative delirium. However, the accuracy of prediction models for patients in the ICU is insufficient when applied to those undergoing cardiovascular surgery²², and delirium prediction models should be constructed only for a specific group of patients²³.

Conventional prediction models based on statistical methods for patients undergoing cardiovascular surgery include those by Koster²⁴ and Rudolph⁴; both reported an area under the receiver operating characteristic curve (AUROC) of 0.75. Although statistical models such as logistic regression are favorable in terms of model interpretability, machine learning is preferred for prediction models²⁵. Clinical prediction models using machine learning algorithms have recently attracted attention. In the area of cardiovascular surgery-associated postoperative delirium, both Mufti²⁶ and Xue²⁷ showed that the prediction performance of machine learning algorithms was superior to that of conventional statistical models, limited to hyperactive delirium and acute kidney injury-related delirium, respectively. Delirium is classified into three subtypes based on the type of psychomotor activity: hyperactive, hypoactive, and mixed¹. Hypoactive delirium reportedly accounts for 92% of all cases of delirium in the cardiovascular ICU²⁸. Therefore, we must assess and deal with delirium, including the hypoactive subtype which is often overlooked in clinical practice.

We aimed to develop and validate new prediction models for cardiovascular surgery-associated postoperative delirium including hypoactive subtype using machine learning algorithms. Early identification of patients at high risk of delirium is critical to effectively implement prevention strategies, such as multicomponent interventions and prophylactic medications.

Methods

Setting and study population

This single-center prospective study included patients aged ≥ 40 years who underwent cardiovascular surgery at Osaka University Hospital between November 2021 and October 2022. The inclusion criteria were patients who underwent coronary artery bypass graft (CABG), valve surgery, ascending aortic replacement (AAR) via a median sternotomy, or minimally invasive cardiac surgery (MICS) using cardiopulmonary bypass (CPB). The exclusion criteria were as follows: (1) patients undergoing transcatheter aortic valve implantation (TAVI) or thoracic endovascular aortic repair (TEVAR); (2) patients managed with deep hypothermic circulatory arrest or selective cerebral perfusion; (3) patients with preoperative delirium; (4) patients diagnosed and treated for dementia preoperatively; (5) patients requiring mechanical ventilation for > 3 days postoperatively, (6) patients with cerebral hemorrhage or stroke within 7 days after surgery; (7) patients requiring reoperation within 7 days after surgery; and (8) patients with suspected alcohol withdrawal delirium. The criteria for alcohol withdrawal delirium were consumption of an average of 60 g or more of alcohol per day immediately prior to admission and the development of delirium within 2 weeks of admission.

Ethics declarations and consent to participate

This study was approved by the Ethical Review Committee of Osaka University Hospital (No: 21158-3; 2021/09/11) and followed the Declaration of Helsinki. All patients provided written informed consent.

Variable selection

The electronic medical records of all patients were prospectively reviewed. We collected 38 preoperative factors, including demographics, medical history, laboratory data, and life history, as well as 10 intraoperative factors, including operation time, pump time, and blood fluid balance. The entire list of 48 variables can be found as Supplementary Table S1 and S2 online. A trained nurse visited the patients on any one day from admission to surgery and assessed their function using the Mini-Cog²⁹ Japanese version, Geriatric Depression Scale-Short Version-Japanese (GDS-S-J)³⁰, and Barthel Index³¹.

Delirium assessment

All patients were followed for 7 days, with day 0 being the day of surgery. Electronic medical records were reviewed daily. The evaluation of delirium began when the patient was extubated in the ICU. A psychiatrist (MH) or a critical nurse (CN) trained by psychiatrists visited patients after discharge from the ICU and when a change in mental status was suspected, then assess delirium according to the Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5)¹. For cases that were difficult to assess, two or more investigators including psychiatrists discussed and came to a final decision. Considering the diurnal variability of delirium, visiting hours were standardized as 3:30 p.m. to 6:30 p.m.

Data preprocessing and statistical analysis

Some values of the intraoperative propofol dose (25.3%) and HbA1c (1.1%) were missing and thus, these factors were excluded from the analysis. The outliers were used after confirming that they were not erroneously entered. Variables are expressed as median (interquartile range [IQR]) or the number of persons (percentage).

Univariate analysis was used to compare patients who did not develop delirium (non-delirium group) and those who did (delirium group). Student's t-test was used for continuous variables normally distributed with equal variance, Welch's t-test for variables normally distributed with unequal variance, Mann-Whitney U test for continuous variables not normally distributed, and Fisher's exact test for categorical variables. We used the Shapiro-Wilk test and Bartlett's test to assess data normality and equality of variance, respectively. All were

two-tailed tests with a significance level of 5%. Corrections for multiple comparisons were not performed because these analyses were exploratory. R software version 4.0.5 was used for statistical processing.

Derivation of prediction models

The features used for classification models were selected based on the results of the univariate analysis and the consultations with two skilled cardiovascular surgeons, three psychiatrists, and a nurse. The process of model development is illustrated in Supplementary Fig. S1 online. We compared the performance of classification models using Bernoulli naive Bayes, Support vector machines³², Random forest³³, Extra-trees³⁴, and XGBoost³⁵. Continuous variables were used as they were. In Bernoulli naive Bayes, continuous variables were binarized to 0 or 1. A binarization threshold is used as a hyperparameter. We searched the hyperparameter spaces of models using a grid search with a fivefold stratified cross-validation. Our dataset is imbalanced because it contains a large proportion of the non-delirium class and a small proportion of the delirium class. Therefore, we used class weights to help the model learn from the imbalanced data. We evaluated the classification performance of these models using a fivefold stratified cross-validation. Seven evaluation metrics were utilized to compare the classification performance of each model: the balanced accuracy, AUROC, area under the precision-recall curve (AUPRC)³⁶, sensitivity, specificity, positive predictive value, and F-value. In addition, a predictive model based on the conventional Logistic Regression (LR) method was also built for comparison with the machine learning models. The LR model was validated using the stratified hold-out method. Our experiments were conducted on Python version 3.8.3.

Result

Patients and occurrence of delirium

Of the 123 patients who met the inclusion criteria and consented to participate in the study, 87 (median age 71 [IQR; 61.5, 75.0] years; 53 [60.9%] males) were included in the analysis. Twenty-four patients (27.6%) developed delirium within seven days after surgery (Fig. 1). Fourteen of the delirious patients assessed delirium during their ICU stay based on electronic medical records, 10 of whom also had delirium symptoms when we visited after discharge from the ICU. No significant difference was found in the length of ICU stay between non-delirium and delirium groups (non-delirium vs. delirium group: median 3 [IQR; 2, 3] days vs. 3 [2, 3], $p=0.976$). The subtypes included 14 (58.3%) hypoactive, eight (33.3%) mixed, including those who did not present with apparent hypoactivity or hyperactivity, and two (8.3%) hyperactive.

Risk factors of delirium

The results of the univariate analysis comparing the non-delirium and delirium groups were shown in Supplementary Table S1 and S2 online. Preoperative factors that showed statistically significant differences were age (non-delirium vs. delirium group: median 69 [IQR; 59.5, 74.0] years vs. 74 [69.0, 79.0], $p=0.004$), use of

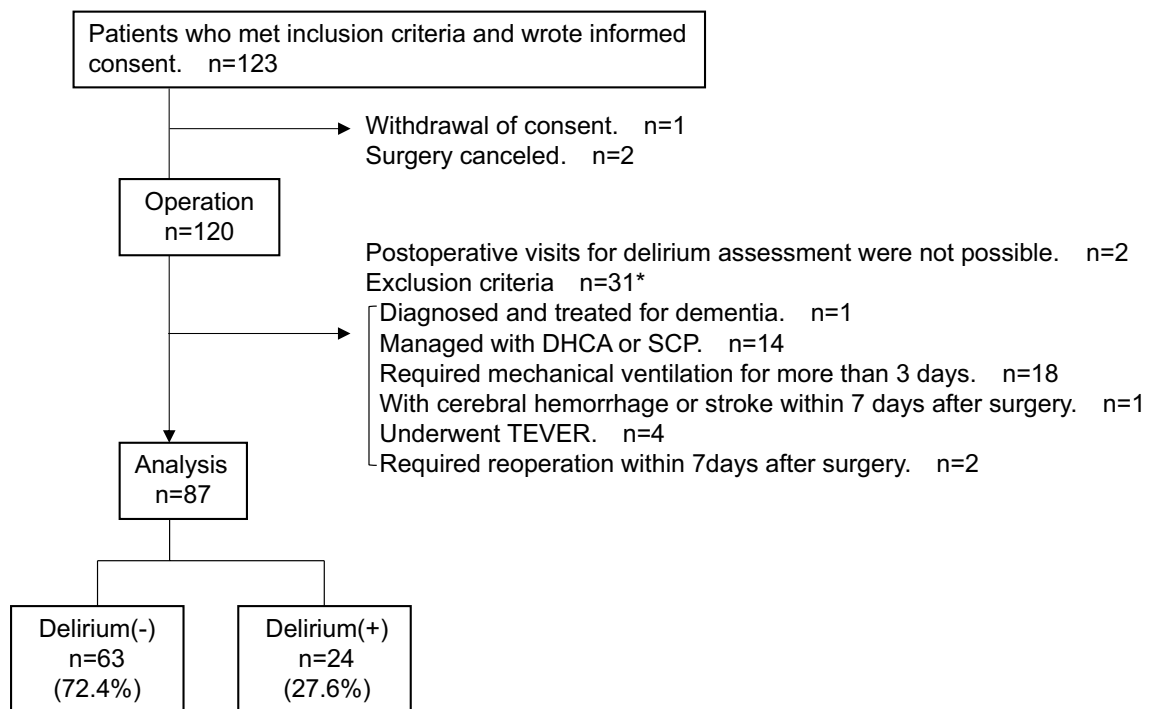


Figure 1. Flow diagram of the study population. DHCA, Deep hypothermic circulatory arrest; SCP, Selective cerebral perfusion; TEVAR, thoracic endovascular aortic repair. *As some patients met more than one exclusion criterion, the total for each criterion does not add up to 31.

psychotropic drugs (including sleeping medications) (6 [9.5%] vs. 9 [37.5%], $p=0.004$), Mini-Cog < 4 (10 [15.9%] vs. 10 [41.7%], $p=0.020$), Barthel Index < 100 (3 [4.8%] vs. 5 [20.8%], $p=0.034$), and CABG (8 [12.7%] vs. 8 [33.3%], $p=0.035$). No intraoperative factors showed statistically significant differences between the groups.

Development of prediction models and validation

In the results of the univariate analysis, CABG, which showed $p < 0.05$, was uninterpretable. First, we removed CABG from the set of features for training the models. Second, we selected a history of stroke or cerebral hemorrhage (non-delirium vs. delirium group, 4 [6.3%] vs. 4 [16.7%], $p=0.208$) and eGFR < 60 (30 [47.6%] vs. 16 [66.7%], $p=0.150$) as the set of features based on the clinical observation. Consequently, we trained all models using the following features: age, use of psychotropic drugs, Mini-Cog < 4, Barthel Index < 100, history of stroke or cerebral hemorrhage, and eGFR < 60.

As a result of cross-validation after hyperparameter tuning, the extra-trees model had the best AUROC (0.76 ± 0.11 [standard deviation]) and AUPRC (0.62 ± 0.18), with a sensitivity of 0.63 and specificity of 0.78. XGBoost showed the best sensitivity (AUROC: 0.75 ± 0.07, AUPRC: 0.59 ± 0.17, sensitivity: 0.67, and specificity: 0.79). The conventional LR model showed lower values for the evaluation metrics than other machine learning models. A comparison of the developed models is shown in Table 1 and the ROC and PR curves are shown in Fig. 2.

Discussion

Herein, we developed models to preoperatively predict up to 67% of patients who develop postoperative delirium based on the following six preoperative factors: age, use of psychotropic drugs, Mini-Cog < 4, Barthel Index < 100, history of stroke or cerebral hemorrhage, and eGFR < 60.

Many studies have tried to predict post- cardiovascular surgery delirium based on statistical methods such as logistic regression^{4,24,37,38}. However, we applied machine learning to predict postoperative delirium, including the hypoactive subtype. Additionally, we placed importance on the prediction of true-positive patients when developing and validating the models using weighting and PR curves³⁶. This resulted in predictive models with similar AUROCs; the sensitivity was more stable than in previous studies. Considering the safety of delirium preventive strategies, such as multicomponent nonpharmacological interventions, a predictive model that emphasizes sensitivity to identify true positives preoperatively may be preferred³⁹.

In this study, no intraoperative factors showed statistically significant differences. Hence, we used only preoperative factors to develop the predictive models. Early prediction of postoperative delirium using only preoperative factors enables us to provide efficient interventions for postoperative delirium, such as premedication, preoperative cognitive training, and adequate staffing. For example, the Extra-trees model we built correctly predicted 63% of true positives, while it indicated that 33% of all participants were not targets of predictive strategies preoperatively. Postoperative delirium is influenced by surgical and anesthesia-related factors, such as the use of anticholinergic drugs including opioids⁴⁰ and CPB, which is associated with inflammatory stimuli-induced nerve injury⁴¹. However, including intraoperative factors in the model reportedly does not improve the prediction accuracy for postoperative delirium⁴². This implies that only preoperative factors may predict the onset of postoperative delirium, which would allow for the initiation of preoperative prophylactic interventions.

We selected features based on the results of univariate analysis and clinical experiments. A prospective observational study reported that Mini-Cog alone showed discriminative power with an AUROC of 0.77 (95%CI; 0.61–0.93) in predicting postoperative delirium, excluding in cardiovascular surgery⁴³. The Mini-Cog is one of the simplest cognitive assessment tools that can be performed in a few minutes²⁹. Psychotropic drugs reportedly have side effects on the central nervous system and are risk factors for delirium⁴⁴. Mini-Cog > 4, use of psychotropic drugs, and a history of stroke or cerebral hemorrhage directly reflect the patient's cerebral function. eGFR < 60 is a standard used to evaluate renal dysfunction, which is known to cause brain dysfunction due to overproduction of inflammatory cytokines and hemodynamic changes^{45–48}. In older patients, microglial immune responses frequently occur in the central nervous system, and excessive neuroinflammation in response to surgical invasion may lead to postoperative delirium and cognitive dysfunction^{49,50}. Some studies cite ADL and instrumental ADL (iADL) impairment as risk factors for delirium^{51,52}, and physical function interventions are recommended for preventing delirium⁵³. Thus, Barthel Index < 100 was included as a risk factor. Immobility due to impairment of physical function can be related to the decreased cholinergic activity, which is important in the pathophysiology of delirium⁵⁴.

Algorithms	AUROC [SD]	AUPRC [SD]	Balanced-accuracy	Sensitivity	Specificity	PPV	F-value
Bernoulli naïve Bayes	0.76 [0.08]	0.59 [0.18]	0.73	0.55	0.94	0.77	0.60
Support vector machine	0.74 [0.08]	0.60 [0.17]	0.74	0.63	0.89	0.68	0.60
Random forest	0.70 [0.15]	0.49 [0.16]	0.68	0.55	0.89	0.56	0.55
Extra-Trees	0.76 [0.11]	0.62 [0.18]	0.70	0.63	0.78	0.52	0.56
XGBoost	0.75 [0.07]	0.59 [0.17]	0.74	0.67	0.79	0.55	0.59
Conventional logistic regression	0.63 [0.15]	0.34 [0.14]	0.68	0.60	0.77	0.50	0.54

Table 1. Comparison of the performance of the developed models. AUROC, area under the receiver operating characteristic curve; SD, standard deviation; AUPRC, area under the precision-recall curve; PPV, positive predictive value.

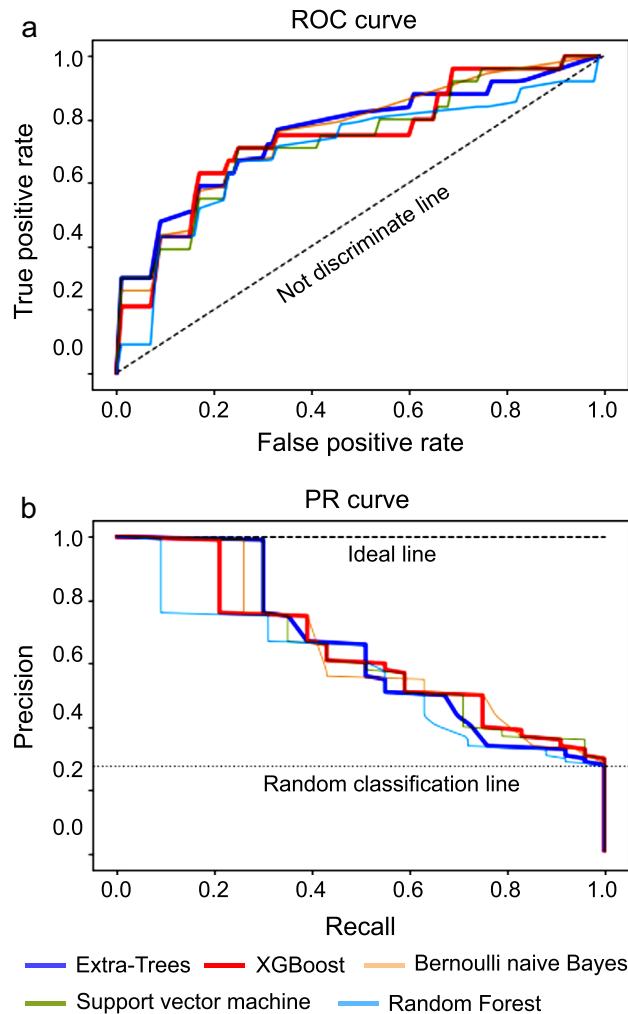


Figure 2. ROC curve and PR curves of the developed models. ROC. curve, Receiver operating characteristic curve; PR curve, Precision-recall curve.

A systematic review in 2021 reported that the most commonly used and best model for predicting delirium in adult inpatients was Random forest, followed by Support vector machine⁵⁵. We used Bernoulli naive Bayes, Support vector machine, Random forest, Extra-trees, and XGBoost and compared the results. Extra-trees and XGBoost, which are superior in terms of AUROC and AUPRC, are ensemble learning models that combine multiple decision trees, similar to that in Random forest. In tree-structured ensemble learning, combining multiple classifiers stabilizes the estimation and suppresses overfitting⁵⁶. Although Random forest did not show particularly good accuracy in this study, such a tree-structured population learning method may be suitable for predicting delirium. The machine learning models we built outperformed conventional LR except for the sensitivity of Random forest and Bernoulli naive Bayes, indicating that machine learning models were better balanced in predicting true positives and true negatives. The advantage of machine learning methods is that they do not have restrictions on their assumption as LR, such as linear distribution of logit transformed values of probability, equivariance of error terms, and the number of predictors⁵⁷. The pathophysiology of delirium is complex and influenced by many factors, machine learning methods, which are highly flexible algorithms, are considered suitable.

Our study has four main limitations. First, some patients with delirium might have been underestimated. Clinical features of delirium are vague, and an estimated 55–70% of hospital delirium cases are missed^{58,59}. The incidence of delirium in this study was 27.6%, which is consistent with a previous report indicating the incidence to be 26–56%^{4–7,60} with prospective diagnosis based on DSM-5 criteria. However, delirium occurring in ICU might have been missed because assessments during ICU stays are based primarily on electronic medical records because ICU access was restricted due to the spread of the novel coronavirus infection.

Second, the subjects and sample size were limited by adopting a prospective design at a single center. Although larger sample sizes are generally preferred for building machine learning models, our study included only 87 participants. In addition, our models were validated using stratified 5-fold cross-validation; no external validation was performed. Our results suggest that machine learning may be able to predict post-cardiovascular surgery delirium with greater sensitivity than conventional statistical methods. The developed models could be integrated

into a meta-model while expanding the sample size in the future to improve their versatility²³. In addition, the sample size affects feature selection. Typical feature selection methods in machine learning models include the wrapper and embedding methods that select variables through model training⁶¹. These methods are used to control model overfitting caused by too many variable dimensions. However, if data used for feature selection is small, there is a risk of causing model overfitting when the method is executed. Thus, unsuitable features for prediction might be selected. Therefore, we applied the statistical feature selection method and discussed it with cardiovascular surgeons, psychiatrists, and a nurse. Machine learning is superior to statistics in terms of discovering data patterns and non-linear relationships⁶², therefore, feature selection methods, such as wrapper and embedding methods, can be applied when we have enough study participants.

Third, the appropriateness of the questionnaire is ambiguous. The mental status of patients aged < 65 years were assessed using the GDS-S-J, which we use in older patients to maintain data consistency. However, the GDS-S-J has been developed specifically for the elderly and thus, its validity in younger adults remains unknown³⁰. A questionnaire developed for all age groups should be applied to assess the mental status more accurately.

Finally, we did not include postoperative factors for prediction. Postoperative factors, such as the duration of ICU stay and ventilator support requirement, reportedly affect postoperative delirium^{63,64}. However, it is also stated that postoperative factors with unclear temporal relationships to delirium onset should not be used for prediction²³. If a significant influence of postoperative factors on the development of delirium is suspected in the target population, taking measures such as revision of the exclusion criteria may be required.

Conclusion

This study's results suggest that applying machine learning algorithms to predict post-cardiovascular surgery delirium shows superior performance, especially in predicting true-positive patients. The developed models were constructed using only preoperative collectible factors, which enables early identification of patients at high risk of delirium. Clinical applications of such predictive models will contribute to the prevention of development of postoperative delirium in patients undergoing cardiovascular surgery by guiding the implementation of preoperative prevention strategies such as non-pharmacological multifactorial interventions and prophylactic medication that focus on high-risk patients.

Data availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 23 May 2023; Accepted: 27 November 2023

Published online: 30 November 2023

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorder*, 5th ed. Washington, DC (2013).
2. Drews, T. *et al.* Postoperative delirium is an independent risk factor for posttraumatic stress disorder in the elderly patient: A prospective observational study. *Eur. J. Anaesthesiol.* **32**, 147–151 (2015).
3. Daiello, L. A. *et al.* Postoperative delirium and postoperative cognitive dysfunction: Overlap and divergence. *Anesthesiology* **131**, 477–491 (2019).
4. Rudolph, J. L. *et al.* Derivation and validation of a preoperative prediction rule for delirium after cardiac surgery. *Circulation* **119**, 229–236 (2009).
5. Schoen, J. *et al.* Preoperative regional cerebral oxygen saturation is a predictor of postoperative delirium in on-pump cardiac surgery patients: A prospective observational trial. *Crit. Care* **15**, R218 (2011).
6. Rudolph, J. L. *et al.* Delirium: An independent predictor of functional decline after cardiac surgery. *J. Am. Geriatr. Soc.* **58**, 643–649 (2010).
7. Brown, C. H. *et al.* The impact of delirium after cardiac surgical procedures on postoperative resource use. *Ann. Thorac. Surg.* **101**, 1663–1669 (2016).
8. Saczynski, J. S. *et al.* Cognitive trajectories after postoperative delirium. *N. Engl. J. Med.* **367**, 30–39 (2012).
9. Koster, S., Hensens, A. G., Schuurmans, M. J. & van der Palen, J. Consequences of delirium after cardiac operations. *Ann. Thorac. Surg.* **93**, 705–711 (2012).
10. Abelha, F. J. *et al.* Outcome and quality of life in patients with postoperative delirium during an ICU stay following major surgery. *Crit. Care* **17**, R257 (2013).
11. Gottesman, R. F. *et al.* Delirium after coronary artery bypass graft surgery and late mortality. *Ann. Neurol.* **67**, 338–344 (2010).
12. Damluji, A. A. *et al.* Older adults in the cardiac intensive care unit: Factoring geriatric syndromes in the management, prognosis, and process of care: A scientific statement from the American Heart Association. *Circulation* **141**, e6–e32 (2020).
13. Aldecoa, C. *et al.* European Society of Anaesthesiology evidence-based and consensus-based guideline on postoperative delirium. *Eur. J. Anaesthesiol.* **34**, 192–214 (2017).
14. Kate Shannon, G. O. D. P. J. S. J. M. C. F. R. N. *American Society for Enhanced Recovery and Perioperative Quality Initiative Joint Consensus Statement on Postoperative Delirium Prevention. Physiology & Behavior* vol. 176 (2016).
15. Inouye, S. K. *et al.* A multicomponent intervention to prevent delirium in hospitalized older patients. *N. Engl. J. Med.* **340**, 669–676 (1999).
16. Oh, E. S., Fong, T. G., Hsieh, T. T. & Inouye, S. K. Delirium in older persons: Advances in diagnosis and treatment. *JAMA* **318**, 1161–1174 (2017).
17. Inouye, S. K., Baker, D. I., Fugal, P. & Bradley, E. H. Dissemination of the hospital elder life program: Implementation, adaptation, and successes. *J. Am. Geriatr. Soc.* **54**, 1492–1499 (2006).
18. Hatta, K. *et al.* Preventive effects of suvorexant on delirium: A randomized placebo-controlled trial. *J. Clin. Psychiatry* **78**, e970–e979 (2017).
19. Hatta, K. *et al.* Preventive effects of ramelteon on delirium: A randomized placebo-controlled trial. *JAMA Psychiatry* **71**, 397–403 (2014).
20. Gosch, M. & Nicholas, J. A. Pharmacologic prevention of postoperative delirium. *Z. Gerontol. Geriatr.* **47**, 105–109 (2014).
21. Curtis, M. S., Forman, N. A., Donovan, A. L. & Whitlock, E. L. Postoperative delirium: Why, what, and how to confront it at your institution. *Curr. Opin. Anaesthesiol.* **33**, 668–673 (2020).

22. Lee, A. *et al.* Risk prediction models for delirium in the intensive care unit after cardiac surgery: A systematic review and independent external validation. *Br. J. Anaesth.* **118**, 391–399 (2017).
23. Lindroth, H. *et al.* Systematic review of prediction models for delirium in the older adult inpatient. *BMJ Open* **8**, e019223 (2018).
24. Koster, S., Hensens, A. G., Schuurmans, M. J. & van der Palen, J. Prediction of delirium after cardiac surgery and the use of a risk checklist. *Eur. J. Cardiovasc. Nurs.* **12**, 284–292 (2013).
25. Boulesteix, A.-L. & Schmid, M. Machine learning versus statistical modeling. *Biom. J.* **56**, 588–593 (2014).
26. Mufti, H. N., Hirsch, G. M., Abidi, S. R. & Abidi, S. S. R. Exploiting machine learning algorithms and methods for the prediction of agitated delirium after cardiac surgery: Models development and validation study. *JMIR Med. Inform.* **7**, e14993 (2019).
27. Xue, X., Chen, W. & Chen, X. A novel radiomics-based machine learning framework for prediction of acute kidney injury-related delirium in patients who underwent cardiovascular surgery. *Comput. Math. Methods Med.* **2022**, 4242069 (2022).
28. McPherson, J. A. *et al.* Delirium in the cardiovascular ICU: Exploring modifiable risk factors. *Crit. Care Med.* **41**, 405–413 (2013).
29. Scanlan, J. & Borson, S. The Mini-Cog: Receiver operating characteristics with expert and naïve raters. *Int. J. Geriatr. Psychiatry* **16**, 216–222 (2001).
30. Sugishita, K. & Asada, T. The creation of the geriatric depression scale-short version-Japanese (GDS-S-J). *Jpn. J. Cognit. Neurosci.* **12**, 87–90 (2009) ((Japanese)).
31. Mahoney, F. I. & Barthel, D. W. Functional evaluation: The Barthel Index. *Md. State Med. J.* **14**, 61–65 (1965).
32. Noble, W. S. What is a support vector machine?. *Nat. Biotechnol.* **24**, 1565–1567 (2006).
33. Breiman, L. E. O. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
34. Geurts, P., Ernst, D. & Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **63**, 3–42 (2006).
35. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* 785–794 (2016).
36. Ozenne, B., Subtil, F. & Maucourt-Boulch, D. The precision–recall curve overcame the optimism of the receiver operating characteristic curve in rare diseases. *J. Clin. Epidemiol.* **68**, 855–859 (2015).
37. de la Varga-Martínez, O. *et al.* Development and validation of a delirium risk prediction preoperative model for cardiac surgery patients (DELIPRECA5): An observational multicentre study. *J. Clin. Anesth.* **69**, 110158 (2021).
38. Segernäs, A. *et al.* Prediction of postoperative delirium after cardiac surgery with a quick test of cognitive speed, mini-mental state examination and hospital anxiety and depression scale. *Clin. Intervent. Aging* **8**, 359–368 (2022).
39. Wong, A. *et al.* Development and validation of an electronic health record-based machine learning model to estimate delirium risk in newly hospitalized patients without known cognitive impairment. *JAMA Netw. Open* **1**, e181018 (2018).
40. Duprey, M. S. *et al.* Opioid use increases the risk of delirium in critically ill adults independently of pain. *Am. J. Respir. Crit. Care Med.* **204**, 566–572 (2021).
41. Saller, T. *et al.* A case series on the value of tau and neurofilament protein levels to predict and detect delirium in cardiac surgery patients. *Biomed. Pap.* **163**, 241–246 (2019).
42. Xue, B. *et al.* Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. *JAMA Netw. Open* **4**, e212240 (2021).
43. Dworkin, A., Lee, D. S. H., An, A. R. & Goodlin, S. J. A simple tool to predict development of delirium after elective surgery. *J. Am. Geriatr. Soc.* **64**, e149–e153 (2016).
44. Kassie, G. M., Nguyen, T. A., Kalisch Ellett, L. M., Pratt, N. L. & Roughead, E. E. Preoperative medication use and postoperative delirium: A systematic review. *BMC Geriatr.* **17**, 298 (2017).
45. Lu, R., Kiernan, M. C., Murray, A., Rosner, M. H. & Ronco, C. Kidney–brain crosstalk in the acute and chronic setting. *Nat. Rev. Nephrol.* **11**, 707–719 (2015).
46. Kinsey, G. R., Li, L. & Okusa, M. D. Inflammation in acute kidney injury. *Nephron. Exp. Nephrol.* **109**, e102–e107 (2008).
47. Liu, M. *et al.* Acute kidney injury leads to inflammation and functional changes in the brain. *J. Am. Soc. Nephrol.* **19**, 1360–1370 (2008).
48. Siew, E. D. *et al.* Acute kidney injury as a risk factor for delirium and coma during critical illness. *Am. J. Respir. Crit. Care Med.* **195**, 1597–1607 (2017).
49. Cunningham, C. Systemic inflammation and delirium: Important co-factors in the progression of dementia. *Biochem. Soc. Trans.* **39**, 945–953 (2011).
50. Li, J., Zong, Y., Cao, X., Tan, L. & Tan, L. Microglial priming in Alzheimer's disease. *Ann. Transl. Med.* **6**, 176–176 (2018).
51. Watt, J. *et al.* Identifying older adults at risk of delirium following elective surgery: A systematic review and meta-analysis. *J. Gen. Intern. Med.* **33**, 500–509 (2018).
52. Han, J. H. *et al.* Delirium in older emergency department patients: Recognition, risk factors, and psychomotor subtypes. *Acad. Emerg. Med.* **16**, 193–200 (2009).
53. Yang, F. M. *et al.* Phenomenological subtypes of delirium in older persons: Patterns, prevalence, and prognosis. *Psychosomatics* **50**, 248–254 (2009).
54. Hshieh, T. T., Fong, T. G., Marcantonio, E. R. & Inouye, S. K. Cholinergic deficiency hypothesis in delirium: A synthesis of current evidence. *J. Gerontol. A Biol. Sci. Med. Sci.* **63**, 764–772 (2008).
55. Chua, S. J., Wrigley, S., Hair, C. & Sahathevan, R. Prediction of delirium using data mining: A systematic review. *J. Clin. Neurosci.* **91**, 288–298 (2021).
56. Dong, X. & Yu, Z. A survey on ensemble learning. *Front. Comput. Sci.* **14**, 241–258 (2020).
57. Christodoulou, E. *et al.* A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* **110**, 12–22 (2019).
58. Hustey, F., Meldon, S. & Palmer, R. Prevalence and documentation of impaired mental status in elderly emergency department patients. *Acad. Emerg. Med.* **7**, 1166 (2000).
59. Inouye, S. K., Foreman, M. D., Mion, L. C., Katz, K. H. & Cooney, L. M. Nurses' recognition of delirium and its symptoms: Comparison of nurse and researcher ratings. *Arch. Intern. Med.* **161**, 2467–2473 (2001).
60. Plaschke, K. *et al.* Early postoperative delirium after open-heart cardiac surgery is associated with decreased bispectral EEG and increased cortisol and interleukin-6. *Intensive Care Med.* **36**, 2081–2089 (2010).
61. Cai, J., Luo, J., Wang, S. & Yang, S. Feature selection in machine learning: A new perspective. *Neurocomputing* **300**, 70–79 (2018).
62. Rose, S. Machine learning for prediction in electronic health data. *JAMA Netw. Open* **1**, e181404 (2018).
63. Chen, H., Mo, L., Hu, H., Ou, Y. & Luo, J. Risk factors of postoperative delirium after cardiac surgery: A meta-analysis. *J. Cardiothorac. Surg.* **16**, 1–12 (2021).
64. Kim, M. Y., Park, U. J., Kim, H. T. & Cho, W. H. DELirium prediction based on hospital information (Delphi) in general surgery patients. *Medicine (Baltimore)* **95**, e3072 (2016).

Acknowledgements

We thank the staff and patients of the Cardiovascular Surgery Unit at the Osaka University Hospital for their cooperation in conducting this study.

Author contributions

C.N., M.H., Y.M., H.M., T.W., and T.U. contributed to the design and conduct of the study and analyzed and interpreted the data. C.N. and T.U. wrote the manuscript and are responsible for the decision to submit it for publication. C.N. contributed to the data collection. C.N. and M.H. assessed the delirium. C.N. and T.K. performed the machine learning analysis. All authors provided critical feedback at various stages of the manuscript and approved its final version.

Funding

This study was supported by JSPS KAKENHI (No: JP22K10839).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-48418-5>.

Correspondence and requests for materials should be addressed to C.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023, corrected publication 2024