Postoperative Delirium Prediction Based on Preoperative Electrocardiogram and Electroencephalogram

Shogo Mito*, Miho Miyajima†, Hirofumi Tomioka†, Hitomi Sato†, Takashi Takeuchi†, Hitoshi Muto†, Yuji Kabasawa†, Hiroyuki Harada†, Kana Eguchi*, Shota Kato*, Manabu Kano*
* Kyoto University, Japan
E-mail: manabu@human.sys.i.kyoto-u.ac.jp
† Tokyo Medical and Dental University, Japan

***Abstract*—Preventing postoperative delirium is of great importance because its onset causes significant social and health problems. We aimed to develop a model to predict the postoperative delirium using data measured preoperatively with electrocardiogram (ECG) sensors and simplified electroencephalogram (EEG) devices in a less invasive situation. The data were collected from patients undergoing oral surgery, specifically from 59 patients (including 8 who developed delirium) from 9 AM until 1 PM and 56 patients (including 8 who developed delirium) from 4 PM until bedtime. We developed binary classification models using support vector machine (SVM) and random forest (RF). We tested three patterns of input variables: EEG indicators, heart rate variability (HRV) indicators, and both. Finally, the SVM model with both indicators was able to predict with a sensitivity of 0.63 and a specificity of 0.69.**

I. Introduction

Delirium is a severe neuropsychiatric syndrome characterized by a sudden decline in cognitive function and attention. Patients often experience altered states of wakefulness, ranging from reduced responsiveness to severe agitation. This condition is frequently accompanied by symptoms such as delusions, hallucinations, and mood changes [1].

The onset of delirium results in numerous adverse effects. In the United States, it is estimated that the extension of hospital stays due to delirium costs an additional 38 billion dollars annually [2]. Furthermore, an increase in mortality rates has been reported, with delirium patients having a mortality rate 2.9 times higher than non-delirium patients [3]. Long-term cognitive decline is also a concern, as the onset of delirium has been reported to be a cause of dementia [4]. Additionally, delirium poses significant psychological distress to patients as well as to their caregivers [5]. Due to the numerous problems caused by delirium, the importance of preventive medicine is emphasized. Predicting the onset of delirium in advance can lead to efficient prevention.

Most predictions of delirium use electronic health records and clinical evaluations by physician's subjectivity [6]. One such model achieved a sensitivity of 72.9% and a specificity of 77.5% using 115 preoperative features [7]. However, clinical evaluations can sometimes differ between trainee doctors and other medical professionals. Hence, highlighting the need for predictions based on objective indicators. Biological markers provide objective measures, and recent studies have successfully predicted delirium with 86% accuracy using preoperative electroencephalogram (EEG) [8]. Additionally, changes in preoperative heart rate variability (HRV) are likely associated with delirium onset [9]. Therefore, this study aims to develop a model for predicting the onset of delirium using these biological indicators.

This work was supported by JSPS KAKENHI Grant Numbers JP20K07915, JP21H03855, and JP23K07006.

II. Methods

This study has been approved by the research ethics committee of Tokyo Medical and Dental University.

In the following text, we defined the period from 9 AM until 1 PM as the AM and the period from 4 PM until bedtime as the PM.

A. Participants

A total of 80 patients aged 16 and older who underwent oral surgery at the Department of Oral and Maxillofacial Surgery, Tokyo Medical and Dental University Hospital, March 23, 2022, and March 25, 2024, were participated in this study. Electrocardiogram (ECG) and EEG were measured twice a day, once in the AM and once in the PM, each for more than 3 minutes, from the day of admission until the day before surgery. We excluded 11 subjects in the AM and 12 subjects in the PM who did not have a Holter ECG attached during EEG measurement. Then, we excluded 7 subjects in the AM and 9 subjects in the PM whose ECG waveforms at the target time were not recorded for some reason. Additionally, Data from 3 subjects in the AM and 3 subjects in the PM, whose ECG waveforms could not be confirmed due to noise, were excluded. As a result, data from 59 subjects in the AM and 56 subjects in the PM were included in the analysis. Tables I and II show the demographic data of the subjects for the AM and PM, respectively.

TABLE I
DEMOGRAPHIC DATA OF STUDY SUBJECTS
(FROM 9 AM UNTIL 1 PM)

Item	Delirium Positive	Delirium Negative
Number of subjects	8	51
Gender (male/female)	5/3	30/21
Age(years±SD)	64.8±12.7	79.3±10.6

TABLE II
DEMOGRAPHIC DATA OF STUDY SUBJECTS
(FROM 4 PM UNTIL BEDTIME)

Item	Delirium Positive	Delirium Negative
Number of subjects	8	48
Gender (male/female)	6/2	27/21
Age(years±SD)	79.8±10.1	63.3±12.9

B. Data Collection

During the measurements, participants were asked to lie down on a bed and close their eyes. ECG was measured at 250 Hz using a Holter monitor Simplo (Technology Inc.). EEG was measured at 128 Hz using a single-channel electroencephalograph ZA-X (Proassist Ltd.). The evaluation of delirium was conducted by physicians over several days postoperatively.

C. EEG Data Preprocessing

We analyzed 3-minute EEG segments with minimal noise that had been preselected by a physician. For each patient, there is one segment in the AM and one segment in the PM. Each segment was divided into 90 non-overlapping epochs, each lasting 2 seconds.

For each epoch, we calculated the kurtosis, skewness, and the difference between the maximum and minimum voltages. For each patient's 3-minutes data segment, we applied the three sigma method to these three metrics to exclude epochs considered to contain noise. Subsequently, for each remaining epoch, we calculated the power ratio for the Theta (4–8 Hz), Alpha (8–13 Hz), Beta (13–30 Hz), and Gamma1 (30–59 Hz) frequency bands, as well as the ratios between these variables, and then computed the average of them over the 3-minutes period.

D. ECG Data Preprocessing

In this study, we analyzed 3-minute segments of ECG taken at the same time as the clipped EEG. We preprocessed the selected data by applying a bandpass filter of 0.5–150 Hz. R-waves were detected by the maximal overlap discrete wavelet transform and RR intervals (RRIs) were calculated from them. During this process, visual inspection was conducted. It is known that accurate HRV indicators cannot be calculated from RRIs containing outliers such as arrhythmias [10]. Hence, we employed the Hampel filter to detect these outliers and replaced them with linearly interpolated values. Subsequently, we calculated the time-domain indicators shown in Table III, the frequency-domain indicators shown in Table IV, and the Poincare plot indicators shown in Table V.

TABLE III
HRV TIME DOMAIN INDICATORS

Feature Name	Definition
meanNN	mean RRI
meanHR	mean heart rate
SDNN	standard deviation of RRI
RMSSD	root mean square of successive RRI differences
SDSD	standard deviation of successive RRI differences
NN50	number of RRI pairs differing by over 50 ms
NN20	number of RRI pairs differing by over 20 ms
pNN50	percentage of RRI pairs differing by over 50 ms
pNN20	percentage of RRI pairs differing by over 20 ms

TABLE IV
HRV FREQUENCY DOMAIN INDICATORS

Feature Name	Definition
HF	Power in the 0.15–0.4 Hz range
LF	Power in the 0.04–0.15 Hz range
LF/HF	ratio of LF to HF

TABLE V
HRV POINCARE PLOT INDICATORS

Feature Name	Definition
SD1	standard deviation of points vertical to identity line
SD2	standard deviation of points horizontal to identity line
SD1*SD2	product of SD1 and SD2
SD2/SD1	ratio of SD2 to SD1

III. MODEL

To distinguish between delirium and non-delirium, we constructed a binary classification model. We designed three types of input variable patterns: EEG indicators, HRV indicators, and both. Considering the circadian rhythm, separate models were built for AM and PM. We used 59 patients (including 8 with delirium) in the AM and 56 patients (including 8 with delirium) in the PM to construct predictive models.

The methods used were support vector machines (SVMs) with a radial basis function (RBF) kernel and random forests (RFs). For hyperparameters, we used the L2 regularization term and RBF kernel parameters for SVMs, and the maximum number of decision branches, the minimum number of samples per leaf node, and the number of input variables randomly selected for each split for RFs. To test classification performance, we employed leave-one-patient-out cross-validation.

After standardizing the training data by setting the mean of each input variable to 0 and the variance to 1, we increased the number of samples for delirium patients to match the number of non-delirium samples using an oversampling method, SMOTE [11]. Oversampling is used to adress significant imbalances between the numbers of samples of minority and majority classes. We performed 10-fold cross-validation on the training data and used Bayesian optimization to determine the hyperparameters that maximize area under the precision-recall curve (PR-AUC). We then made predictions on the remaining one sample and validated the performance. Finally, we calculated SHAP values [12] for each model to determine the importance of the variables.

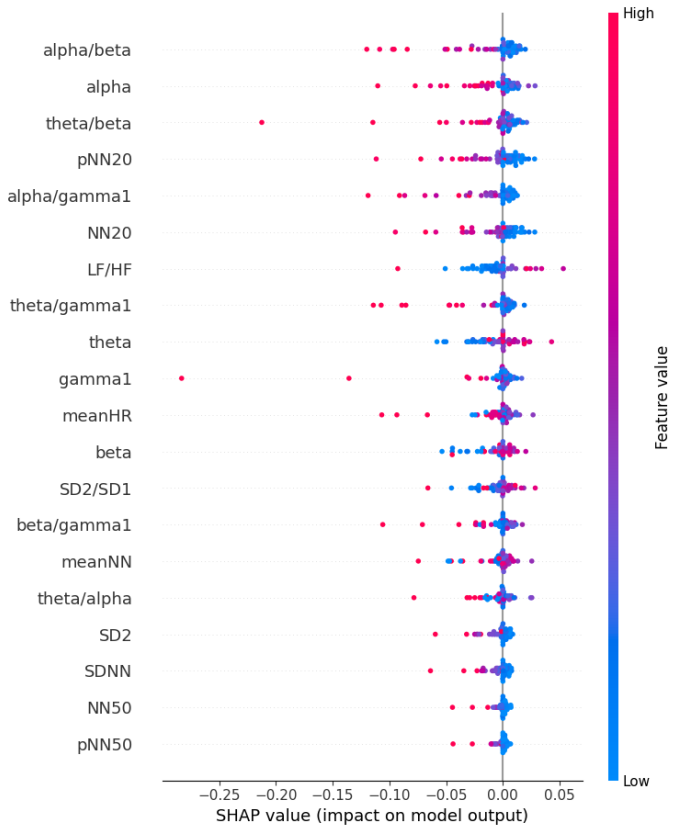

Fig. 1. Shap Value on SVM Model (from 9 AM until 1 PM)

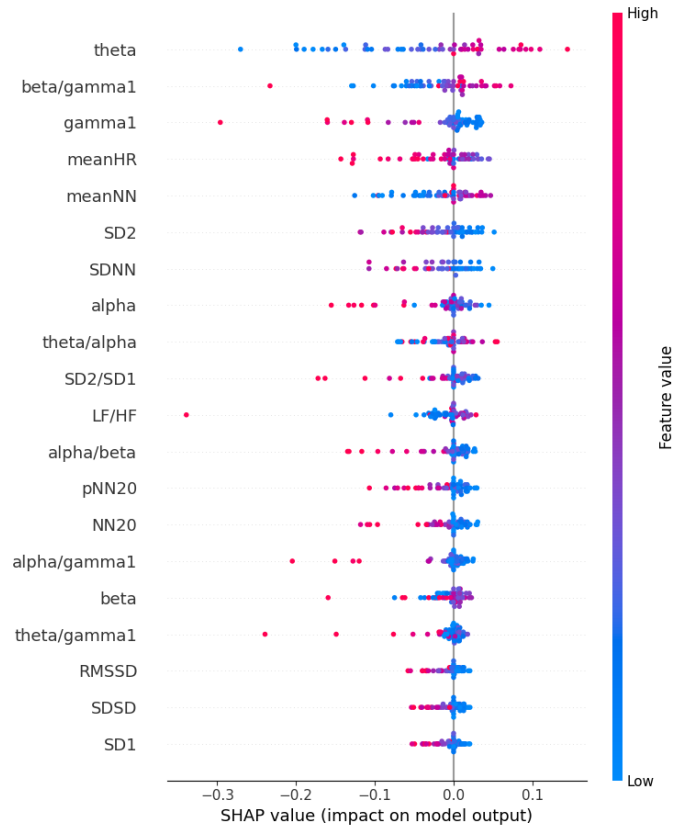

Fig. 2. Shap Value on SVM Model (from 4 PM until bedtime)

IV. Results

Tables VI and VII show the results in the AM and the PM, respectively, indicating the average accuracy, sensitivity, specificity, and PR-AUC on the training data. Tables VIII and IX summarize the prediction results for the test data in the AM and PM, respectively. Changing the input variable pattern did not lead to significant changes in performance.

Figs. 1 and 2 plot the SHAP values for all test data predicted by the SVM model with the highest sensitivity, using both EEG and HRV indicators in the AM and PM, respectively. The SHAP values are sorted in descending order of their mean absolute values. The horizontal axis represents the SHAP values, where larger values indicate a positive contribution to the model and smaller values indicate a negative contribution. The relative magnitude of SHAP values for different variables indicates their comparative importance in influencing the model's output. The redder plot shows the value of the respective feature is higher. The bluer plot shows the value of the respective feature is lower.

V. Discussion

Since sensitivity is important for predicting delirium, we focus on sensitivity in the test data. We conclude that the SVM model using both EEG and HRV indicators or only EEG indicators for AM, and the Random Forest model using only HRV indicators for PM, performed best in our data set. However, compared to the previous study, which reported an accuracy of 86% [8], the accuracy in our study is lower. This discrepancy may be attributed to the fact that 30% of the patients in the previous study were delirium patients, while only 12.5% of the patients in our study were delirium patients, resulting in greater class imbalance. Additionally, while the previous study used three EEG channels, our study only used one EEG channel, potentially failing to capture the EEG changes associated with the onset of delirium.

Next, We discuss the top-ranked variables in importance. First, we will discuss the EEG indicators. The low power of alpha waves contributes to delirium, which is consistent with the study showing that preoperative EEG alpha waves are inversely correlated with the severity of postoperative delirium [13]. The increase in theta wave power contributes to the onset of delirium, consistent with research indicating that delirium patients experience delayed brain waves and increased theta waves [14], suggesting that delayed brain waves may occur preoperatively. The increase in gamma wave activity indicates a heightened state of arousal [15], implying that preoperative arousal levels may be lower in delirium patients compared to non-delirium patients.

Finally, we will discuss the indicators related to HRV. The decrease in SD2 suggesting the onset of delirium is consistent with the study indicating that SD2 decreases under stress conditions [16]. It is noted that higher SDNN indicates greater

TABLE VI
TRAIN AVERAGE RESULTS (FROM 9 AM UNTIL 1 PM)

Method	Input Variable	Accuracy	Sensitivity	Specificity	PR-AUC
SVM	HRV	0.84	0.83	0.86	0.92
	EEG	0.82	0.92	0.72	0.94
	HRV and EEG	0.81	0.90	0.71	0.90
RF	HRV	0.85	0.80	0.90	0.96
	EEG	0.89	0.86	0.89	0.96
	HRV and EEG	0.84	0.84	0.85	0.93

TABLE VII
TRAIN AVERAGE RESULTS (FROM 4 PM UNTIL BEDTIME)

Method	Input Variable	Accuracy	Sensitivity	Specificity	PR-AUC
SVM	HRV	0.88	0.91	0.85	0.96
	EEG	0.85	0.86	0.84	0.88
	HRV and EEG	0.86	0.99	0.73	0.96
RF	HRV	0.86	0.80	0.90	0.96
	EEG	0.83	0.92	0.74	0.93
	HRV and EEG	0.89	0.92	0.86	0.98

TABLE VIII
TEST RESULTS (FROM 9 AM UNTIL 1 PM).
FRACTIONS IN PARENTHESES INDICATE HOW MANY PATIENTS ANSWERED CORRECTLY

Method	Input Variable	Accuracy	Sensitivity	Specificity	PR-AUC
SVM	HRV	0.81	0.50 (4/8)	0.86 (44/51)	0.20
	EEG	0.66	0.63(5/8)	0.67(34/51)	0.16
	HRV and EEG	0.68	0.63(5/8)	0.69(35/51)	0.20
RF	HRV	0.73	0.38(3/8)	0.78(40/51)	0.18
	EEG	0.73	0.50(4/8)	0.76(39/51)	0.21
	HRV and EEG	0.68	0.38(3/8)	0.69(36/51)	0.12

TABLE IX
TEST RESULTS (FROM 4 PM UNTIL BEDTIME).
FRACTIONS IN PARENTHESES INDICATE HOW MANY PATIENTS ANSWERED CORRECTLY

Method	Input Variable	Accuracy	Sensitivity	Specificity	PR-AUC
SVM	HRV	0.75	0.38(3/8)	0.81(39/48)	0.14
	EEG	0.75	0.38(3/8)	0.81(39/48)	0.23
	HRV and EEG	0.64	0.63(5/8)	0.65(31/48)	0.18
RF	HRV	0.68	0.63(5/8)	0.69(33/48)	0.24
	EEG	0.61	0.38(3/8)	0.65(31/48)	0.25
	HRV and EEG	0.64	0.50(4/8)	0.69(33/48)	0.18

resistance to stress [17], and lower SDNN is associated with a state of stress, aligning with the prediction that a decrease in SDNN contributes to the onset of delirium. Additionally, when SDNN decreases, the variability of NN intervals diminishes, leading to reductions in NN20 and pNN20. Therefore, the decrease in NN20 and pNN20 due to stress load is consistent with their contribution to the onset of delirium. A higher LF/HF ratio contributes to delirium, indicating a state of mental stress [18], consistent with studies suggesting that stress is partly responsible for the onset of delirium.

VI. CONCLUSIONS

In this study, our model was able to predict with a sensitivity of 0.63, a specificity of 0.69, and an accuracy of 0.68. However, due to the limited number of our data set, we need to conduct further studies. Although our models did not achieve the accuracy reported in other studies, the result allowed us to investigate the relationship between physiological markers and the delirium. In the future work, we plan to collect additional subject data and compare the results with those obtained using various machine learning techniques.

REFERENCES

[1] J. E. Wilson, M. F. Mart, C. Cunningham, *et al.*, "Delirium," *Nature Reviews Disease Primers*, vol. 6, p. 90, Nov. 2020. DOI: 10.1038/s41572-020-00223-4.

[2] D. L. Leslie, E. R. Marcantonio, Y. Zhang, L. Leo-Summers, and S. K. Inouye, "One-year health care costs associated with delirium in the elderly population," *Archives of Internal Medicine*, vol. 168, pp. 27–32, Jan. 2008. DOI: 10.1001/archinternmed.2007.4.

[3] D. K. Kiely, E. R. Marcantonio, S. K. Inouye, *et al.*, "Persistent delirium predicts greater mortality," *Journal of the American Geriatrics Society*, vol. 57, pp. 55–61, Jan. 2009. DOI: 10.1111/j.1532-5415.2008.02092.x.

[4] T. E. Goldberg, C. Chen, Y. Wang, *et al.*, "Association of delirium with long-term cognitive decline: A meta-analysis," *JAMA Neurology*, vol. 77, pp. 1373–1381, Nov. 2020. DOI: 10.1001/jamaneurol.2020.2273.

[5] S. T. Williams, J. K. Dhesi, and J. S. L. Partridge, "Distress in delirium: Causes, assessment and management," *European Geriatric Medicine*, vol. 11, pp. 63–70, Feb. 2020. DOI: 10.1007/s41999-019-00276-z. eprint: 2019Dec9.

[6] T. Strating, L. Shafiee Hanjani, I. Tornvall, R. Hubbard, and I. A. Scott, "Navigating the machine learning pipeline: A scoping review of inpatient delirium prediction models," *BMJ Health Care Informatics*, vol. 30, e100767, Jul. 2023. DOI: 10.1136/bmjhci-2023-100767.

[7] A. Bishara, C. Chiu, E. L. Whitlock, *et al.*, "Postoperative delirium prediction using machine learning models and preoperative electronic health record data," *BMC Anesthesiology*, vol. 22, p. 8, Jan. 2022. DOI: 10.1186/s12871-021-01543-y.

[8] M. Hata, Y. Miyazaki, C. Nagata, *et al.*, "Predicting postoperative delirium after cardiovascular surgeries from preoperative portable electroencephalography oscillations," *Frontiers in Psychiatry*, vol. 14, p. 1 287 607, Nov. 2023. DOI: 10.3389/fpsyt.2023.1287607.

[9] G. Ernst, L. O. Watne, M. Rostrup, and B. E. Neerland, "Delirium in patients with hip fracture is associated with increased heart rate variability," *Aging Clinical and Experimental Research*, vol. 32, pp. 2311–2318, Nov. 2020. DOI: 10.1007/s40520-019-01447-5. eprint: 2020Jan8.

[10] "Heart rate variability. standards of measurement, physiological interpretation, and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology," en, *Eur. Heart J.*, vol. 17, no. 3, pp. 354–381, Mar. 1996.

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002, ISSN: 1076-9757. DOI: 10.1613/jair.953. [Online]. Available: http://dx.doi.org/10.1613/jair.953.

[12] S. Lundberg and S.-I. Lee, *A unified approach to interpreting model predictions*, 2017. arXiv: 1705.07874 [cs.AI]. [Online]. Available: https://arxiv.org/abs/1705.07874.

[13] L. Acker, M. K. Wong, M. C. Wright, *et al.*, "Preoperative electroencephalographic alpha-power changes with eyes opening are associated with postoperative attention impairment and inattention-related delirium severity," *British Journal of Anaesthesia*, vol. 132, pp. 154–163, Jan. 2024. DOI: 10.1016/j.bja.2023.10.037. eprint: 2023Dec12.

[14] C. S. Guay, M. Kafashan, E. R. Huels, *et al.*, "Postoperative delirium severity and recovery correlate with electroencephalogram spectral features," *Anesthesia Analgesia*, vol. 136, pp. 140–151, Jan. 2023. DOI: 10.1213/ANE.0000000000006075. eprint: 2022May13.

[15] C. Summerfield, A. I. Jack, and A. P. Burgess, "Induced gamma activity is associated with conscious awareness of pattern masked nouns," *International Journal of Psychophysiology*, vol. 44, pp. 93–100, May 2002. DOI: 10.1016/s0167-8760(02)00003-x.

[16] S. Rahman, M. Habel, and R. J. Contrada, "Poincaré plot indices as measures of sympathetic cardiac regulation: Responses to psychological stress and associations with pre-ejection period," *International Journal of Psychophysiology*, vol. 133, pp. 79–90, Nov. 2018. DOI: 10.1016/j.ijpsycho.2018.08.005. eprint: 2018Aug11.

[17] H. G. Kim, E. J. Cheon, D. S. Bai, Y. H. Lee, and B. H. Koo, "Stress and heart rate variability: A meta-analysis and review of the literature," *Psychiatry Investigation*, vol. 15, pp. 235–245, Mar. 2018. DOI: 10.30773/pi.2017.08.17. eprint: 2018Feb28.

[18] W. von Rosenberg, T. Chanwimalueang, T. Adjei, U. Jaffer, V. Goverdovsky, and D. P. Mandic, "Resolving ambiguities in the lf/hf ratio: Lf-hf scatter plots for the categorization of mental and physical stress from hrv," *Frontiers in Physiology*, vol. 8, p. 360, Jun. 2017. DOI: 10.3389/fphys.2017.00360.