Journal of Clinical Anesthesia 102 (2025) 111771

Contents lists available at ScienceDirect

Journal of Clinical Anesthesia

journal homepage: www.elsevier.com/locate/jclinane

Original Contribution

Functional MRI-based machine learning strategy for prediction of postoperative delirium in cardiac surgery patients: A secondary analysis of a prospective observational study

Mei-Yan Zhou [a,b,1], Yi-Bing Shi [c,1], Sheng-Jie Bai [c,1], Yao Lu [d], Yan Zhang [d], Wei Zhang [e], Wei Wang [e], Yang-Zi Zhu [b,d,*], Jun-Li Cao [d,f,**], Li-Wei Wang [a,b,***]

[a] Department of Anesthesiology, Xuzhou Clinical College of Xuzhou Medical University, Xuzhou, Jiangsu 221009, China
[b] Department of Anesthesiology, Xuzhou Central Hospital, Xuzhou, Jiangsu 221009, China
[c] Department of Medical Imaging, Xuzhou Central Hospital, Xuzhou, Jiangsu 221009, China
[d] Jiangsu Province Key Laboratory of Anesthesiology, Xuzhou Medical University, Xuzhou, Jiangsu 221004, China
[e] Department of Cardiac Surgery, Xuzhou Central Hospital, Xuzhou, Jiangsu 221009, China
[f] Department of Anesthesiology, The Affiliated Hospital of Xuzhou Medical University, Xuzhou Medical University, Xuzhou, Jiangsu 221006, China

HIGHLIGHTS

- This study utilized fMRI data to construct an artificial intelligence model to predict postoperative delirium.
- The trained random forest model exhibited excellent performance in predicting postoperative delirium.
- The most discriminative nodes for prediction were located in the default, cingulo-opercular, and frontoparietal networks.

ARTICLE INFO

Keywords:
Postoperative delirium
Functional magnetic resonance imaging
Machine learning
Prediction model
Cardiac surgery

ABSTRACT

Study objective: Delirium is a common complication after cardiac surgery and is associated with poor prognosis. An effective delirium prediction model could identify high-risk patients who might benefit from targeted prevention strategies. We introduce machine learning models that employ resting-state functional MRI datasets obtained before surgery to predict postoperative delirium.
Design: A secondary analysis of a prospective observational study.
Setting: The study was conducted at one tertiary hospital in China.
Patients: The study involved 103 patients who underwent preoperative functional MRI scan and cardiac valve replacement.
Interventions: None.
Measurements: Delirium was assessed twice daily for the first seven postoperative days using the Confusion Assessment Method. We used three whole-brain functional connectivity (FC) measures (parcel-wise connectivity matrix, mean FC and degree of FC) and trained three machine models, namely, random forest, logistic regression, and linear support vector machine, to distinguish delirium patients from patients without delirium. The top performing model was selected for further training with functional MRI datasets and clinical variables.
Main results: This study included 103 participants. A total of 29 participants (28.2 %) met postoperative delirium criteria. Based solely on functional MRI datasets, the random forest model trained using the degree of FC achieved the highest accuracy (0.864), precision (0.887), specificity (0.894), F1 score (0.859) and area under the curve (0.924), and this model was further optimized for accuracy (0.879), sensitivity (0.909), F1 score (0.882) and area under the curve (0.928) by fusing clinical variables. The most discriminative nodes for predicting postoperative delirium were located in the default, cingulo-opercular, and frontoparietal networks.

* Correspondence to: Y.-Z. Zhu, Department of Anesthesiology, Xuzhou Central Hospital, Xuzhou, Jiangsu 221009, China.
** Correspondence to: J.-L. Cao, Jiangsu Province Key Laboratory of Anesthesiology, Xuzhou Medical University, Xuzhou, Jiangsu 221004, China.
*** Correspondence to: L.-W. Wang, Department of Anesthesiology, Xuzhou Clinical College of Xuzhou Medical University, Xuzhou, Jiangsu 221009, China.
E-mail addresses: zhuyz@xzhmu.edu.cn (Y.-Z. Zhu), caojl0310@aliyun.com (J.-L. Cao), doctorlww@sina.com (L.-W. Wang).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.jclinane.2025.111771
Received 20 May 2024; Received in revised form 27 December 2024; Accepted 31 January 2025
Available online 13 February 2025

Conclusions: This study found that the random forest model using preoperative functional MRI data and clinical variables was accurate in identifying patients at high risk of developing delirium after cardiac surgery.

1. Introduction

Delirium is a syndrome characterized by acutely occurred and fluctuating disturbances of cognition, attention, and consciousness [1,2]. Postoperative delirium (POD) is a common complication of surgery and is associated with poor short- and long-term sequelae, including nosocomial complications, increased hospital stay, worsening cognitive decline, increased mortality, and loss of independence [1,3]. The incidence of POD varies with the type of surgery and patient age, occurring in up to 54.9 % of older patients after cardiac surgery [4]. Although POD is often transient, it is an indicator of brain insufficiency and a harbinger of worsened outcomes.

The underlying mechanisms behind delirium have not been fully elucidated. Several hypotheses of delirium pathology include neuroinflammation, neurotransmitter interference and functional disconnection [5–8]. However, there is currently no effective way to treat delirium or eliminate its adverse effects, highlighting the importance of preventing this disease [9]. Studies on effective prevention strategies have made some progress in recent years [10–13]. Early accurate prediction of postoperative delirium can help medical staff target high-risk patients to prevent delirium [14].

Despite research identifying several risk factors or prediction models for delirium based on clinical data [15,16], there is no validated preoperative prediction rule for postoperative delirium in complex clinical situations. Functional MRI provides a reliable method for identifying brain dysfunction based on the correlation of blood oxygen level-dependent (BOLD) signal fluctuations between different brain regions. Our previous study revealed perioperative changes in brain functional connectivity in cardiac surgery patients using resting-state fMRI [17]. Similar findings were observed for other types of surgery [18,19]. However, no studies have examined the prediction of postoperative delirium based on preoperative functional connectivity features.

Machine learning methods based on data-driven approaches to neuroimaging are increasingly utilized to overcome the problem of prediction or treatment selection for individuals suffering from a variety of neuropsychiatric disorders [20–23]. Studies have shown that graph theory and machine learning approaches can be utilized to predict the progression of mild cognitive impairment to Alzheimer's disease in patients using resting-state fMRI [24,25]. In addition, addiction treatment completion was well predicted using machine learning pattern classification of fMRI data in persons with substance abuse problems [26]. In this study, we extracted connectivity (correlation) maps from pairs of brain regions in a whole-brain analysis before cardiac surgery. We subsequently trained machine learning models to identify features that distinguished patients with postoperative delirium. We anticipate that this study will establish a method for accurately predicting postoperative delirium incidence preoperatively, thereby enabling more effective identification of high-risk patients.

2. Materials and methods

2.1. Participants

This was a second analysis of database from a prospective observational study registered in the Chinese Clinical Trial Registry (ChiCTR-OOC-17012542) [17]. During the underlying study, we enrolled patients who underwent valve replacement surgery under cardiopulmonary bypass (CPB), had an education level above sixth grade, and a Mini-Mental State Examination (MMSE) score $\geq$ 23. Patients with a history of craniocerebral surgery, cerebrovascular disease, hepatorenal failure, psychiatric illness, alcoholism, illiteracy, left-handedness, or metal implants incompatible with MRI were excluded. In the current study, patients over 60 years of age who completed a preoperative MRI scan were selected from the previously enrolled population. The study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of Xuzhou Central Hospital (XZXY-LK-20240823-0128).

2.2. Perioperative management

Perioperative management was performed as described in our previous study [17]. All patients underwent general anesthesia according to our standardized protocol. Anesthesia was induced using midazolam (0.05 mg/kg), cisatracurium (0.3 mg/kg), etomidate (0.3 mg/kg), and sufentanil (5 μg/kg), and maintained with remifentanil, sevoflurane, and propofol, with the Bispectral Index (BIS) maintained between 40 and 60. Heart rate, arterial pressure, respiratory rate, body temperature, $PETCO_2$, and SpO_2 were continuously monitored. All patients underwent surgery with standard CPB. Nasopharyngeal or rectal temperature during CPB was maintained under mild hypothermia (32 °C). Intraoperative blood salvage and α-stat pH management were employed. Perfusion pressure was maintained between 60 and 80 mmHg with norepinephrine, while pump flow was kept at 2.0–2.5 L/min/m^2. Hematocrit was kept above 21 % during CPB and above 25 % throughout the remaining perioperative period. The body rewarming rate was maintained at approximately 0.25 °C per minute. All patients received standardized postoperative pain management with hydromorphone.

2.3. Clinical assessment and delirium diagnosis

All participants underwent baseline measurements and clinical assessments performed by trained research staff. Age, education, gender, diabetes, hypertension, body mass index, left ventricular ejection fraction, Mini-Mental State Examination score, procedure duration, CPB pump duration, and aortic cross-clamp time were extracted from patient medical records. Delirium was defined according to the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders. Patients were assessed continuously for 7 days after surgery twice daily. The delirium assessment was performed by trained research staff using the Confusion Assessment Method (CAM) and CAM for the Intensive Care Unit (CAM-ICU). Measurements were conducted in the morning between 8:00 AM and 12:00 PM, and in the afternoon between 4:00 PM and 8:00 PM [27].

2.4. MRI acquisition

Patients underwent MRI scans (Siemens Skyra 3 Tesla scanner with a 20-channel head coil) before surgery. High-resolution sagittal three-dimensional magnetization-prepared rapid acquisition with gradient echo structural images was acquired with the following parameters: matrix, 256 × 256; field-of-view, 256 mm × 224 mm; 192 one-millimetre-thick slices; echo time, 2.98 ms; repetition time, 2530 ms. Axial T2-weighted imaging images were acquired with the following parameters: matrix, 320 × 320; field-of-view, 230 mm × 230 mm; 18 six-millimetre-thick slices; echo time, 99 ms; and repetition time, 6000 ms. Rs-fMRI data were acquired with the following parameters: matrix, 64 × 64; field-of-view, 220 mm × 220 mm; 35 three-millimetre-thick slices; echo time, 30 ms; and repetition time, 2000 ms.

2.5. MRI data preprocessing

Resting-state fMRI data were preprocessed using SPM12 (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, London, UK). The first ten time points were discarded to avoid magnetic coil saturation. The remaining images were corrected for slice timing and subsequently realigned. Next, the T1 images were co-registered with the realigned images and segmented into gray matter, white matter, and cerebrospinal fluid. The functional images were then spatially normalized into Montreal Neurological Institute (MNI) space using transformations from segmentation, resampled to $3 \times 3 \times 3$ mm^3 voxels, and smoothed with a 6-mm full-width at half-maximum (FWHM) isotropic Gaussian kernel. After preprocessing, bandpass filtering (0.008–0.09 Hz), detrending, and regression of six motion parameters and their first-order derivatives, along with signals from white matter and cerebrospinal fluid (using the CompCor strategy), were further applied.

2.6. Functional connectivity analysis

The functional connectivity matrix of each subject was calculated using the Dosenbach 160 (DOS160) atlas in MATLAB (R2020a) [28]. Specifically, blood oxygen level-dependent (BOLD) time courses were extracted for each ROI. Subsequently, Pearson correlation coefficients between these fMRI time series of each ROI pair were computed to construct an adjacency matrix. Fisher-*Z*-transformation was applied to this adjacency matrix, ensuring that the resulting elements of the final functional connectivity (FC) matrix for each subject followed a normal distribution. A subject-specific threshold was used to extract the top 10 % correlation values from the matrix, thereby enhancing the fidelity of FC assessments both intra- and interindividually. Additionally, the mean FC for each ROI, quantified as the mean of its nonzero correlation coefficients, was computed. The FC degree was quantified as the ratio of the number of connections (i.e., nonzero values in the upper decile matrix) to the aggregate number of feasible connections per parcel grounded in graph theory.

2.7. Classification models and statistical analysis

In the present study, 29 individuals with postoperative delirium were identified as the positive class, labelled '1', while 74 individuals without postoperative delirium were designated the negative class, labelled '0'. The features included functional connectivity between brain regions calculated using DOS160, mean FC, and degree FC.

Machine learning was implemented using nested cross-validation, with the outer layer consisting of stratified tenfold cross-validation (Fig. 1). The synthetic minority oversampling technique (SMOTE) was utilized for oversampling to balance the sample sizes of the positive and negative classes. The normalization of each fold of the outer layer involved the use of a z score, feature selection, and hyperparameter tuning of the model via the training data. Feature selection was conducted using the F score method. The F score is defined as follows:

$$F(i) \equiv \frac{\left(\overline{x}_i^{(+)} - \overline{x}_i\right)^2 + \left(\overline{x}_i^{(-)} - \overline{x}_i\right)^2}{\frac{1}{n_+-1}\sum_{k=1}^{n_+}\left(x_{k,i}^{(+)} - \overline{x}_i^{(+)}\right)^2 + \frac{1}{n_--1}\sum_{k=1}^{n_-}\left(x_{k,i}^{(-)} - \overline{x}_i^{(-)}\right)^2}$$

where $\overline{x}_i$, $\overline{x}_i^{(+)}$, and $\overline{x}_i^{(-)}$ are the averages of the kth features of the whole, positive, and negative datasets, respectively; $x_{k,i}^{(+)}$ is the ith feature of the kth positive instance; and $x_{k,i}^{(-)}$ is the ith feature of the kth negative instance. The numerator represents the distinction between the positive and negative groups, while the denominator reflects the distinction within each group. A higher F score suggests a greater discriminative capability of the feature.

For mean FC and degree FC, the top 10 features were retained; for the functional connectivity matrix, the top 60 features were retained. When combining optimal features with clinical characteristics, due to the inclusion of several clinical features (age, gender, education, body mass index, MMSE score, diabetes, hypertension, and left ventricular ejection fraction), which have been previously associated with POD risk [4,16,29], the number of retained features in the original degree model was adjusted from 10 to 20. Hyperparameter selection was carried out using inner-layer cross-validation, which in this case was also stratified by 10-fold cross-validation. The support vector machine (SVM) and logistic regression models underwent hyperparameter optimization for parameter C, with a range of [2–5,25], encompassing 50 uniformly distributed values on a logarithmic scale. For the random forest algorithm, hyperparameter optimization was conducted for n_estimators, ranging from 50 to 200 with increments of 10. The optimization-prioritized accuracy was used as the standard. Following the completion of the inner-layer cross-validation optimization, the optimal hyperparameters were used to train the training set of the outer-layer cross-validation. Subsequently, standardization and feature selection were applied to the test set of the outer-layer cross-validation, and predictions were made using the model. In the feature selection process within the nested cross-validation framework, special attention was given to features that were consistently selected in every fold of the outer loop. These features are typically considered robust and predictive across varying subsets of the data. To effectively communicate the results of this feature selection process, we proceeded with the

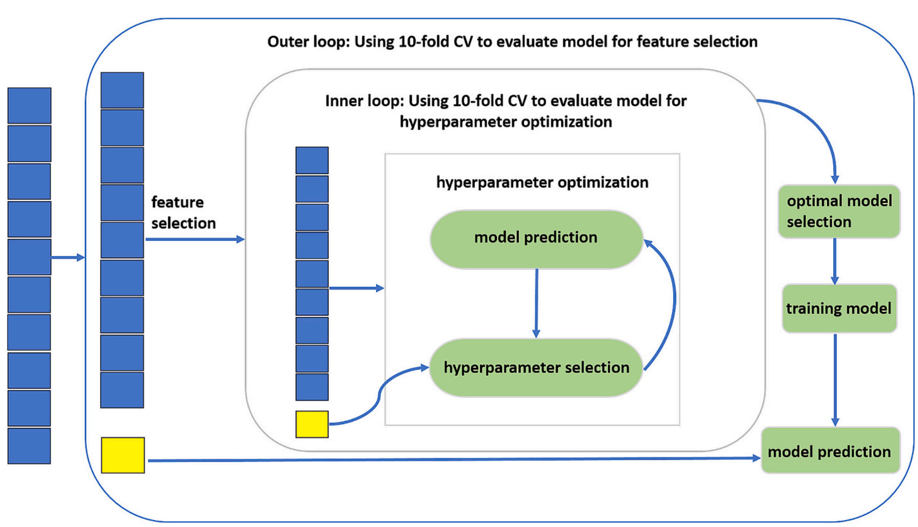

Fig. 1. The process of nested cross-validation of machine learning models.

visualization and detailed reporting of the identified features.

Finally, the results from all the instances of the outer-layer cross-validation in the machine learning process were aggregated for the computation of machine learning evaluation metrics. The labels were permuted 5000 times, and the process was repeated to obtain empirical *p* values for each evaluation metric. A p value less than 0.05 was considered to indicate statistical significance.

3. Results

3.1. Characteristics of the study population

In total, 103 patients with sufficient quality preoperative fMRI scans and postoperative delirium assessments were included. Twenty-nine patients (28.2 %) developed delirium within the first seven post-operative days. No significant differences were observed in the basic demographic or clinical characteristics between the no-delirium group and the delirium group (Table 1).

3.2. Discriminative brain regions based on features of functional connectivity

Using models trained on the degree of FC for POD versus no-POD classification, the most discriminative regions, which included the anterior cingulate, med cerebellum, temporal cortex, mid insula, dorsal frontal cortex, and sup frontal cortex, were demonstrated as key features in each of the three models. It depicts the weight or importance of the degree of FC for the 6 nodes (Table S1, Fig. S1). Discriminative brain regions based on features of the mean FC and connectivity matrix in the LR and SVM models are shown (Figs. S2 and S3, Tables S2 and S3). As noted earlier, the RF model based on the degree of FC along with clinical variables achieved the best performance, and the discriminative brain regions, such as the ACC and dFC, based on this model were similar to those described above (Fig. 2, Table S4).

3.3. Performance of predictive models for delirium

The predictive performance results for POD based on the three models and three FC features are summarized (Table S5). The accuracies of the three machine learning models were 0.811 (RF), 0.682 (LR) and 0.712 (SVM) using mean FC. The accuracies of the three machine learning models were 0.864 (RF), 0.811 (LR) and 0.811 (SVM) according to the degree of FC. The accuracies of the three machine learning models were 0.818 (RF), 0.803 (LR) and 0.803 (SVM) using a connectivity matrix. In addition to the accuracy, the RF model trained using the degree of FC also achieved the highest precision, specificity, F1 score and area under the curve (Fig. 3A, B, C).

Considering the contribution of clinical characteristics to previous POD prediction models, we compared the prediction performance of the RF model based only on clinical variables, including age, gender, education, body mass index, MMSE score, diabetes, hypertension and left ventricular ejection fraction, and the RF model based on the degree of FC and clinical variables. The performance of the degree of functional connectivity (FC) alone outperformed that of clinical variables alone on every indicator of the random forest (RF) model. The performance of the RF model based on the degree of FC and clinical variables was better (accuracy: 0.879; AUC: 0.928) than using these two types of data alone (Fig. 3D, Table 2). We also used 11 features (Table S4) selected by the RF model, based on the degree of functional connectivity and clinical variables, to build a risk score for the prediction of POD. A total score of 6 or higher indicated delirium with an accuracy of 0.886 (Table S6).

Table 1
Patient demographics and clinical and surgical characteristics.

Characteristics	POD (*N* = 29)	Non-POD (*N* = 74)	P value
Age, mean (SD), y	68.0 (5.1)	65.7 (6.0)	0.07
Education, median (IQR), y	9.1 (8–10)	9.9 (8–11.75)	0.08
Female, n (%)	20 (69.0)	40 (54.1)	0.17
BMI, mean (SD), kg/m^2	24.8 (1.0)	25.1 (1.5)	0.41
Diabetes, n (%)	6 (20.7)	6 (8.1)	0.07
Hypertension, n (%)	6 (20.7)	9 (13.8)	0.27
MMSE, median (IQR), point	26.2 (25–27)	26.0 (25–27)	0.58
LVEF, mean (SD), %	53.4 (3.7)	54.7 (3.7)	0.11
ASA class, n (%)			
2	5	11	0.76
3	24	63	
Duration of procedure, mean (SD), min	199.1 (23.5)	197.4 (27.9)	0.77
Duration of pump CPB, mean (SD), min	104.1 (14.1)	104.2 (17.9)	0.99
Aortic cross clamp time, mean (SD), min	75.0 (9.4)	74.1 (10.3)	0.67

Abbreviations: BMI, body mass index; CPB, cardiopulmonary bypass; SD, standard deviation; IQR, interquartile range; LVEF, left ventricular ejection fraction; MMSE, Mini-Mental State Examination, min minute; POD, postoperative delirium; y, year.

4. Discussion

In the present study, we found that data-driven machine learning models could be used to reliably predict POD using features derived from preoperative resting-state fMRI. The RF model based on the degree of FC and clinical variables achieved the highest accuracy and AUC. The most discriminative brain regions for predicting delirium encompassed the anterior cingulate, dorsal frontal cortex, precuneus/posterior cingulate, angular gyrus, med cerebellum, and temporal lobe, most of which are critical nodes in the frontoparietal network (FPN), the default mode network (DMN) and the cingulo-opercular network (CON).

We applied three machine learning models (RF, LR, and SVM) to classify POD, and each model was trained on three whole-brain FC measures: the mean FC, degree of FC and connectivity matrix. The best predictive model for discriminating POD patients from non-POD patients was the RF model based on the degree of FC, with an accuracy of 86.4 % and an AUC of 0.924. Prior to this work, other prediction models for POD were usually based on clinical characteristics, such as age, sex, education, diabetes status, hypertension status, and type of surgery [29]. The reported AUCs ranged from 0.54 to 0.90 in prediction models of POD based on clinical characteristics [29]. Here, we identified the great advantages of fMRI data for machine learning predictive models.

Although resting-state fMRI BOLD signals have strong stability for disease diagnosis [30], it is also important to choose suitable metrics from fMRI data. A previous study showed that preoperative global connectivity strength was not predictive of POD development in elderly people undergoing different types of surgery [31]. In the present study, compared with those of the mean FC and connectivity matrix, the three machine learning models based on the degree of FC achieved optimal performance. This may be because the degree of FC is more representative of each node's role in the higher-level FC topology of the brain [32,33]. Considering that fMRI provides a large amount of 4-dimensional data, model selection is also important. We employed three machine learning methods on the training dataset to fit the parameters and construct the respective predictive models. Compared with logistic regression and support vector machines, a random forest model based on the mean FC, degree of FC and connectivity matrix achieved optimal performance. This was expected because random forests generally provide high predictive accuracy, especially when dealing with complex and high-dimensional data [34]. Overall, the machine learning approach includes the entire process, from extracting the most relevant features of the functional brain connectome to model cross-validation, and can provide reliable predictions regarding POD.

Notably, the value of using clinical characteristics to predict POD cannot be ignored, although clinical characteristics are highly variable. A previous study showed that training models on combined fMRI and

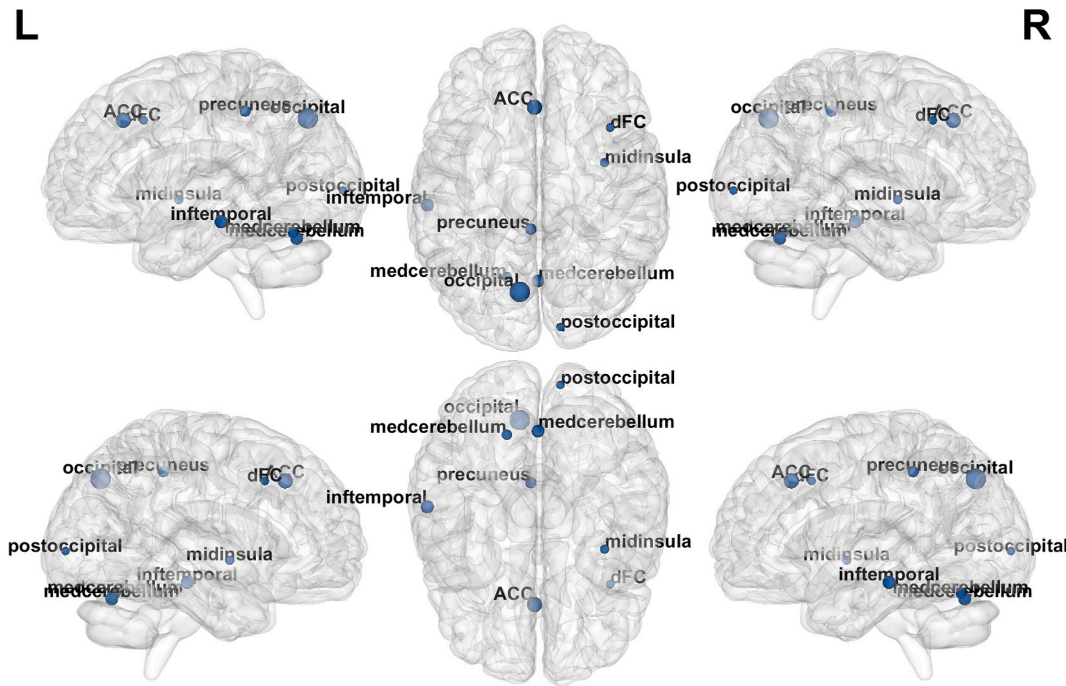

Fig. 2. Discriminative brain regions from the random forest (RF) model based on the degree of functional connectivity (FC) and clinical variables. The size of the node represents the average weight of each node. Abbreviations: ACC, anterior cingulate cortex; dFC, dorsal frontal cortex.

clinical features can improve patient compliance prediction performance [26]. We found that the performance of the RF model based on the degree of FC and clinical characteristics was slightly better than that based only on the degree of FC in terms of accuracy, sensitivity, F1 Score and AUC, but relatively poor in precision and specificity. These results indicate that clinical characteristics have both value and variability in the prediction of POD.

The aims of machine learning modelling tend to not only focus on the model's prediction ability but also may help to elucidate mechanisms underlying susceptibility to POD in the context of brain functional networks. The current study used machine learning methods to determine how each brain region contributes to identify patients at high risk for POD based on feature weights or importance. We have found that several important nodes are located in the frontoparietal network and default mode network. Consistent with prior research, these networks have been recognized as essential components underpinning human cognition [35–37]. Our previous research revealed that the functional connectivity of nodes between the frontoparietal network and default mode network decreased after surgery but was indeed enhanced before surgery [17]. An EEG-based study had similar findings: increased frontal alpha band connectivity before delirium [7]. These changes are thought to act as a functional compensatory mechanism to maintain cognitive function. Yet the characteristic changes in functional connectivity before surgery were not very substantial, and it seemed difficult to determine vulnerability to POD solely by functional network characteristics [38]. However, our study showed that machine learning models, such as the cingulo-opercular network, can optimize this process and identify previously unreported network nodes for POD. The cingulo-opercular network is critical for action and physiological control, arousal, alertness, errors, and pain [39–41], and its functions overlap strongly with the clinical manifestations of delirium. Overall, these findings suggest that the default mode network, the executive control network, and the cingulo-opercular network may be involved in susceptibility to delirium.

There were several limitations in the current study. First, all participants underwent valve replacement surgery with cardiopulmonary bypass, leading to a possible lack of generalizability to other populations. Considering the high incidence of POD after cardiac surgery and the pathophysiological differences among different operations, our focus on cardiac surgery patients guarantees the validity of the prediction model but may also limit its application. Second, the clinical variables included in this study are limited. Including additional clinical variables related to POD, such as geriatric depression scale scores, alcohol use, and sleep disorders, could further enhance the model's performance. In addition, models are built on a finite number of data points in a single medical centre. In future studies, we will pool data from different medical centres to increase the number of training samples and repeat the validation of the prediction model in an independent sample. Finally, it is important to recognize that fMRI scanning is time-consuming and expensive compared to other imaging modalities, which may limit the model's clinical application. The use of structural MRI data or the development of novel scanning techniques, such as ultrafast brain MRI [42], may help address the time and cost constraints in future research.

5. Conclusions

In the current study, we found that employing a machine learning approach with preoperative resting-state fMRI data and clinical variables provides high classification accuracy for POD in cardiac surgery patients and captures neuroimaging features of the brain functional connectome in POD patients. This research provided a potential method for identifying cardiac surgery patients undergoing cardiopulmonary bypass who are at high risk of POD and may facilitate the early implementation of preventive measures.

Funding

This study was supported in part by grants from the National Natural Science Foundation of China (NSFC82171912 to YZZ; NSFC82293641 to J-LC; NSFC82071903 to LWW); the National Key Research and Development Program of China—Sci-Tech Innovation 2030-Major Project (2021ZD0203100 to JLC); the Jiangsu Medical Key Discipline/Laboratory Construction Project (JSDW202231 to LWW) and the Xuzhou Pengcheng Talent Project (XWRCHT20220017 to MYZ).

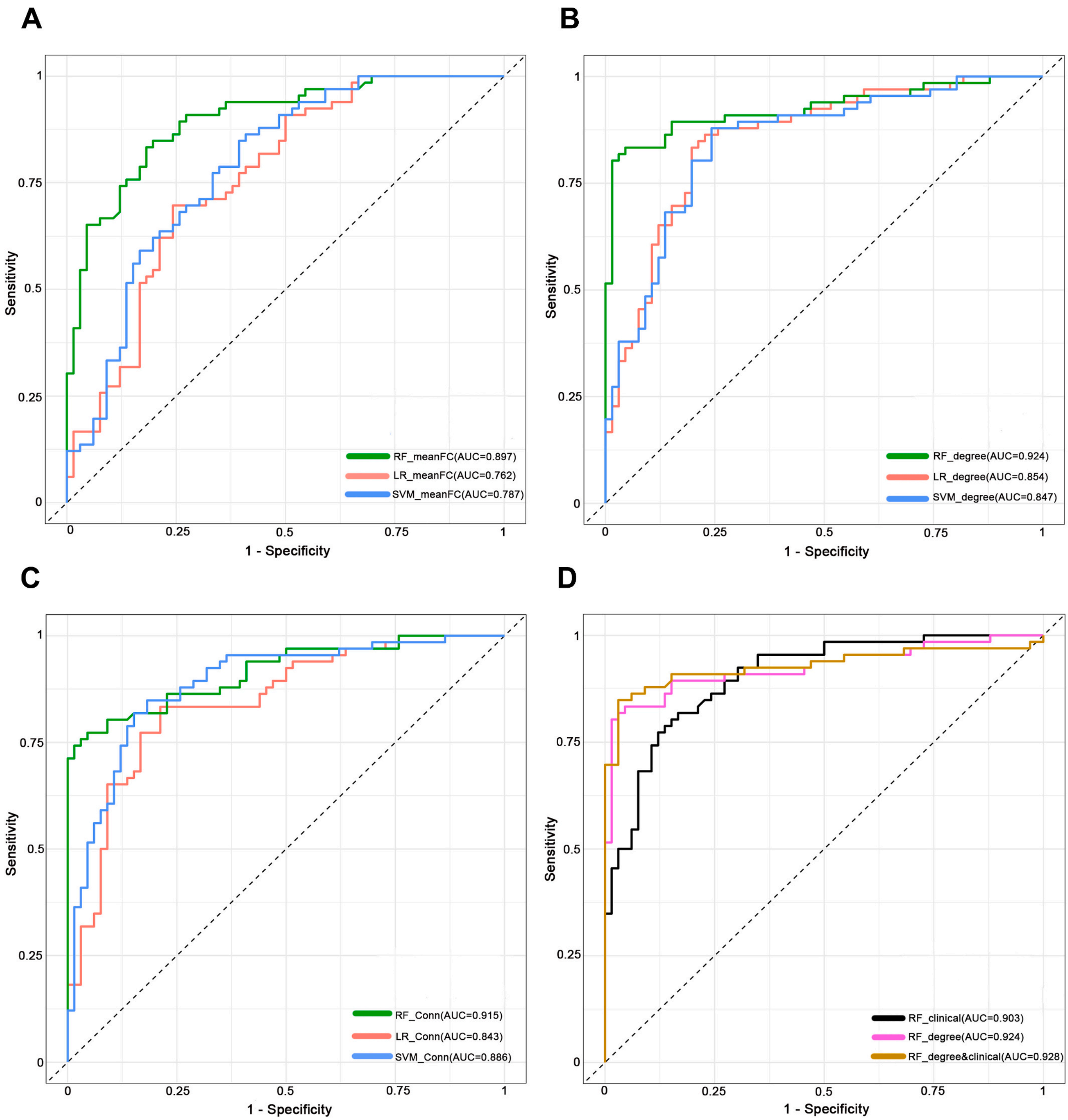

Fig. 3. Performances of machine learning models for predicting POD. (A) The ROC curves of three models based on mean FC, including a random forest (RF), a logistic regression (LR) and a linear support vector machine (SVM). (B) The ROC curves of the three models based on the degree of FC. (C) The ROC curves of the three models based on the connectivity matrix. (D) The ROC curves of the RF model based on the degree of FC and clinical variables. For each ROC curve, the AUC is indicated in the respective legend.

Table 2
Performance of the random forest model based on the clinical variables and degree of FC.

Measures	Accuracy	Precision	Sensitivity	Specificity	F1.score	AUC
Clinical variables	0.803	0.794	0.818	0.788	0.806	0.903
Degree of FC	0.864	0.887	0.833	0.894	0.859	0.924
Clinical variables & degree of FC	0.879	0.857	0.909	0.848	0.882	0.928

Abbreviations: AUC, area under the curve; FC, functional connectivity.

Authors contribution

YZZ, LWW and JLC initiated, designed and supervised the research. MYZ, YBS, SJB, YL, YZ, WZ and WW performed the study and acquired data. MYZ, YBS, SJB, and YZZ analyzed data. YZZ and MYZ wrote the manuscript. All authors read and approved the manuscript.

CRediT authorship contribution statement

Mei-Yan Zhou: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Yi-Bing Shi:** Writing – review & editing, Software, Resources, Methodology, Formal analysis, Data curation. **Sheng-Jie Bai:** Writing – review & editing, Software, Resources, Methodology, Formal analysis. **Yao Lu:** Software, Resources, Formal analysis, Data curation. **Yan Zhang:** Software, Resources, Formal analysis, Data curation. **Wei Zhang:** Software, Funding acquisition, Formal analysis. **Wei Wang:** Resources, Methodology, Formal analysis. **Yang-Zi Zhu:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Jun-Li Cao:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Li-Wei Wang:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We acknowledge the contributions of the statistical guidance of Yanmei Liu, PhD, from the China Evidence-based Medicine Center, West China Hospital, Sichuan University. We also would like to thank Jicheng Qiu, SN ENGR, for technical support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinane.2025.111771.

Data availability

The data sets used for the current study are available from the corresponding authors upon reasonable request.

References

[1] Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. Lancet 2014;383:911–22. https://doi.org/10.1016/s0140-6736(13)60688-1.
[2] Mattison MLP. Delirium. Ann Intern Med 2020;173:Itc49–itc64. https://doi.org/10.7326/aitc202010060.
[3] Jin Z, Hu J, Ma D. Postoperative delirium: perioperative assessment, risk reduction, and management. Br J Anaesth 2020;125:492–504. https://doi.org/10.1016/j.bja.2020.06.063.
[4] Smulter N, Lingehall HC, Gustafson Y, Olofsson B, Engström KG. Delirium after cardiac surgery: incidence and risk factors. Interact Cardiovasc Thorac Surg 2013;17:790–6. https://doi.org/10.1093/icvts/ivt323.
[5] Han Y, Zhang W, Liu J, Song Y, Liu T, Li Z, et al. Metabolomic and lipidomic profiling of preoperative CSF in elderly hip fracture patients with postoperative delirium. Front Aging Neurosci 2020;12:570210. https://doi.org/10.3389/fnagi.2020.570210.
[6] Sumner M, Deng C, Evered L, Frampton C, Leslie K, Short T, et al. Processed electroencephalography-guided general anaesthesia to reduce postoperative delirium: a systematic review and meta-analysis. Br J Anaesth 2023;130:e243–53. https://doi.org/10.1016/j.bja.2022.01.006.
[7] Tanabe S, Mohanty R, Lindroth H, Casey C, Ballweg T, Farahbakhsh Z, et al. Cohort study into the neural correlates of postoperative delirium: the role of connectivity and slow-wave activity. Br J Anaesth 2020;125:55–66. https://doi.org/10.1016/j.bja.2020.02.027.
[8] Taylor J, Parker M, Casey CP, Tanabe S, Kunkel D, Rivera C, et al. Postoperative delirium and changes in the blood-brain barrier, neuroinflammation, and cerebrospinal fluid lactate: a prospective cohort study. Br J Anaesth 2022;129:219–30. https://doi.org/10.1016/j.bja.2022.01.005.
[9] Groen JA, Banayan D, Gupta S, Xu S, Bhalerao S. Treatment of delirium following cardiac surgery. J Card Surg 2012;27:589–93. https://doi.org/10.1111/j.1540-8191.2012.01508.x.
[10] Cao SJ, Zhang Y, Zhang YX, Zhao W, Pan LH, Sun XD, et al. Delirium in older patients given propofol or sevoflurane anaesthesia for major cancer surgery: a multicentre randomised trial. Br J Anaesth 2023;131:253–65. https://doi.org/10.1016/j.bja.2023.04.024.
[11] Enomoto K, Kosaka S, Kimura T, Matsubara M, Kitada Y, Mieno M, et al. Prevention of postoperative delirium after cardiovascular surgery: a team-based approach. J Thorac Cardiovasc Surg 2023;165:1873–81. e2, https://doi.org/10.1016/j.jtcvs.2021.07.027.
[12] Li YW, Li HJ, Li HJ, Zhao BJ, Guo XY, Feng Y, et al. Delirium in older patients after combined epidural-general anesthesia or general anesthesia for major surgery: a randomized trial. Anesthesiology 2021;135:218–32. https://doi.org/10.1097/aln.0000000000003834.
[13] Tao M, Zhang S, Han Y, Li C, Wei Q, Chen D, et al. Efficacy of transcranial direct current stimulation on postoperative delirium in elderly patients undergoing lower limb major arthroplasty: a randomized controlled trial. Brain Stimul 2023;16:88–96. https://doi.org/10.1016/j.brs.2023.01.839.
[14] Marcantonio ER, Goldman L, Mangione CM, Ludwig LE, Muraca B, Haslauer CM, et al. A clinical prediction rule for delirium after elective noncardiac surgery. Jama 1994;271:134–9.
[15] Chen H, Mo L, Hu H, Ou Y, Luo J. Risk factors of postoperative delirium after cardiac surgery: a meta-analysis. J Cardiothorac Surg 2021;16:113. https://doi.org/10.1186/s13019-021-01496-w.
[16] Rudolph JL, Jones RN, Levkoff SE, Rockett C, Inouye SK, Sellke FW, et al. Derivation and validation of a preoperative prediction rule for delirium after cardiac surgery. Circulation 2009;119:229–36. https://doi.org/10.1161/circulationaha.108.795260.
[17] Zhu Y, Zhou M, Jia X, Zhang W, Shi Y, Bai S, et al. Inflammation disrupts the brain network of executive function after cardiac surgery. Ann Surg 2023;277:e689–98. https://doi.org/10.1097/sla.0000000000005041.
[18] Huang H, Tanner J, Parvataneni H, Rice M, Horgas A, Ding M, et al. Impact of total knee arthroplasty with general anesthesia on brain networks: cognitive efficiency and ventricular volume predict functional connectivity decline in older adults. J Alzheimers Dis 2018;62:319–33. https://doi.org/10.3233/jad-170496.
[19] Yang S, Xiao W, Wu H, Liu Y, Feng S, Lu J, et al. Management based on multimodal brain monitoring may improve functional connectivity and post-operative neurocognition in elderly patients undergoing spinal surgery. Front Aging Neurosci 2021;13:705287. https://doi.org/10.3389/fnagi.2021.705287.
[20] Baecker L, Garcia-Dias R, Vieira S, Scarpazza C, Mechelli A. Machine learning for brain age prediction: introduction to methods and clinical applications. EBioMedicine 2021;72:103600. https://doi.org/10.1016/j.ebiom.2021.103600.
[21] Boutet A, Madhavan R, Elias GJB, Joel SE, Gramer R, Ranjan M, et al. Predicting optimal deep brain stimulation parameters for Parkinson's disease using functional MRI and machine learning. Nat Commun 2021;12:3043. https://doi.org/10.1038/s41467-021-23311-9.
[22] Saccà V, Sarica A, Novellino F, Barone S, Tallarico T, Filippelli E, et al. Evaluation of machine learning algorithms performance for the prediction of early multiple sclerosis from resting-state FMRI connectivity data. Brain Imaging Behav 2019;13:1103–14. https://doi.org/10.1007/s11682-018-9926-9.
[23] van Boven MR, Henke CE, Leemhuis AG, Hoogendoorn M, van Kaam AH, Königs M, et al. Machine learning prediction models for neurodevelopmental outcome after preterm birth: a scoping review and new machine learning evaluation framework. Pediatrics 2022:150. https://doi.org/10.1542/peds.2021-056052.
[24] Hojjati SH, Ebrahimzadeh A, Khazaee A, Babajani-Feremi A. Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM. J Neurosci Methods 2017;282:69–80. https://doi.org/10.1016/j.jneumeth.2017.03.006.
[25] Hojjati SH, Ebrahimzadeh A, Khazaee A, Babajani-Feremi A. Predicting conversion from MCI to AD by integrating rs-fMRI and structural MRI. Comput Biol Med 2018;102:30–9. https://doi.org/10.1016/j.compbiomed.2018.09.004.
[26] Steele VR, Maurer JM, Arbabshirani MR, Claus ED, Fink BC, Rao V, et al. Machine learning of functional magnetic resonance imaging network connectivity predicts substance abuse treatment completion. Biol Psychiatry Cogn Neurosci Neuroimaging 2018;3:141–9. https://doi.org/10.1016/j.bpsc.2017.07.003.
[27] Huet O, Gargadennec T, Oilleau JF, Rozec B, Nesseler N, Bouglé A, et al. Prevention of post-operative delirium using an overnight infusion of dexmedetomidine in patients undergoing cardiac surgery: a pragmatic, randomized, double-blind, placebo-controlled trial. Crit Care 2024;28:64. https://doi.org/10.1186/s13054-024-04842-1.
[28] Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, et al. Prediction of individual brain maturity using fMRI. Science 2010;329:1358–61. https://doi.org/10.1126/science.1194144.
[29] Cai S, Li J, Gao J, Pan W, Zhang Y. Prediction models for postoperative delirium after cardiac surgery: systematic review and critical appraisal. Int J Nurs Stud 2022;136:104340. https://doi.org/10.1016/j.ijnurstu.2022.104340.
[30] Ingabire HN, Qu H, Li M, He S, Amos JT, Cui Y, et al. Stability analysis of fMRI BOLD signals for disease diagnosis. IEEE Trans Neural Syst Rehabil Eng 2022;30:967–78. https://doi.org/10.1109/tnsre.2022.3164074.

[31] Ditzel FL, van Montfort SJT, Vernooij LM, Kant IMJ, Aarts E, Spies CD, et al. Functional brain network and trail making test changes following major surgery and postoperative delirium: a prospective, multicentre, observational cohort study. Br J Anaesth 2023;130:e281–8. https://doi.org/10.1016/j.bja.2022.07.054.
[32] Bassett DS, Sporns O. Network neuroscience. Nat Neurosci 2017;20:353–64. https://doi.org/10.1038/nn.4502.
[33] van den Heuvel MP, Sporns O. Rich-club organization of the human connectome. J Neurosci 2011;31:15775–86. https://doi.org/10.1523/jneurosci.3539-11.2011.
[34] Capitaine L, Genuer R, Thiébaut R. Random forests for high-dimensional longitudinal data. Stat Methods Med Res 2021;30:166–84. https://doi.org/10.1177/0962280220946080.
[35] DeSerisy M, Ramphal B, Pagliaccio D, Raffanello E, Tau G, Marsh R, et al. Frontoparietal and default mode network connectivity varies with age and intelligence. Dev Cogn Neurosci 2021;48:100928. https://doi.org/10.1016/j.dcn.2021.100928.
[36] Marek S, Dosenbach NUF. The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. Dialogues Clin Neurosci 2018;20:133–40. https://doi.org/10.31887/DCNS.2018.20.2/smarek.
[37] Smallwood J, Bernhardt BC, Leech R, Bzdok D, Jefferies E, Margulies DS. The default mode network in cognition: a topographical perspective. Nat Rev Neurosci 2021;22:503–13. https://doi.org/10.1038/s41583-021-00474-4.
[38] Matar E, Sanders RD. Where there's smoke': longitudinal cognitive disintegration after postoperative delirium? Br J Anaesth 2023;131:1–5. https://doi.org/10.1016/j.bja.2023.02.025.
[39] Coste CP, Kleinschmidt A. Cingulo-opercular network activity maintains alertness. Neuroimage 2016;128:264–72. https://doi.org/10.1016/j.neuroimage.2016.01.026.
[40] Gordon EM, Chauvin RJ, Van AN, Rajesh A, Nielsen A, Newbold DJ, et al. A somato-cognitive action network alternates with effector regions in motor cortex. Nature 2023;617:351–9. https://doi.org/10.1038/s41586-023-05964-2.
[41] Pierce JE, Petro NM, Clancy E, Gratton C, Petersen SE, Neta M. Specialized late cingulo-opercular network activation elucidates the mechanisms underlying decisions about ambiguity. Neuroimage 2023;279:120314. https://doi.org/10.1016/j.neuroimage.2023.120314.
[42] Schuhholz M, Ruff C, Bürkle E, Feiweier T, Clifford B, Kowarik M, et al. Ultrafast brain MRI at 3 T for MS: evaluation of a 51-second deep learning-enhanced T2-EPI-FLAIR sequence. Diagnostics (Basel) 2024:14. https://doi.org/10.3390/diagnostics14171841.