Journal of Anesthesia (2025) 39:355–365
https://doi.org/10.1007/s00540-025-03471-4

ORIGINAL ARTICLE

Association between intraoperative electroencephalograph complexity index and postoperative delirium in elderly patients undergoing orthopedic surgery: a prospective cohort study

Xiao-yi Hu[1] · Yu-chen Dai[2] · Lan-yue Zhu[2] · Jian-jun Yang[3] · Jie Sun[2] · Mu-huo Ji[1]

Received: 13 October 2024 / Accepted: 15 February 2025 / Published online: 4 March 2025

Abstract

Purpose The primary method for predicting POD (postoperative confusion) relies on the analysis of clinical features. Brain activity complexity is a promising factor associated with the state of consciousness. The aim of this study was to investigate the role of EEG (electroencephalography) complexity changes in predicting POD in elderly patients undergoing orthopedic surgery.

Methods From January 2024 to August 2024, 289 elderly patients undergoing orthopedic surgery were recruited at the Second Affiliated Hospital of Nanjing Medical University. Intraoperative EEG data from patients were collected and then EEG nonlinear features were extracted by MATLAB custom scripts. The logistic regression and CNN (convolutional neural networks) were used to explore the predictive effect of nonlinear features on POD from both static and dynamic perspectives.

Results Low permutation Lempel–Ziv complexity (PLZC) among the EEG nonlinear features emerged as an independent risk factor for POD [OR = 0.210; 95% CI (0.050–0.850); $p = 0.029$]. Receiver operating characteristic curve (ROC) analysis revealed a poor area under the curve of 0.615 (95% CI 0.517–0.711) for PLZC in predicting POD. After the inclusion of temporal factors, the ROC analysis indicated that the EEG nonlinear indices had a moderate predictive effect on POD [AUC = 0.701; (95% CI 0.541–0.862)].

Conclusions EEG nonlinear feature indices may be effective biomarkers for POD and could help predict POD in elderly patients undergoing orthopedic surgery.

Keywords Electroencephalogram · Complexity · Postoperative delirium · EEG

Introduction

Postoperative delirium (POD), recognized as a sudden disruption in attention and cognitive processes, is especially prevalent and severe among elderly patients [1]. The emergence of POD can trigger a series of adverse effects, leading to a decline in autonomy, increased morbidity and mortality, as well as escalated healthcare costs [2]. The significant influence of POD on functional capabilities and life quality has extensive social consequences, affecting not only individuals but also families, communities, and the healthcare system as a whole [3]. To date, due to the complexity of POD's pathogenesis, specific methods to prevent its occurrence are still undiscovered, and diagnosis primarily relies on clinical observation [4]. Therefore, understanding the underlying mechanisms of POD and developing effective preventative and therapeutic measures are crucial for the advancement of medical care and the overall wellbeing of society.

Xiao-yi Hu and Yu-chen Dai have contributed equally to this work.

✉ Jie Sun
dgsunjie@hotmail.com

✉ Mu-huo Ji
jimuhuo2019@126.com

[1] Department of Anesthesiology, The Second Affiliated Hospital of Nanjing Medical University, Nanjing, China

[2] Department of Anesthesiology, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China

[3] Department of Anesthesiology, Pain and Perioperative Medicine, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

At present, the predominant approach to predicting POD relies on the analysis of clinical characteristics [5]. However, this approach is potentially limited due to the apparent non-linear relationships between the variables and outcomes, suggesting a need for more advanced analytical [6, 7]. Electroencephalography (EEG) serves as a non-invasive, rapid, and effective diagnostic tool for assessing global neuronal activity in the brain [8]. Its utility, however, is challenged by the inherent difficulty in interpreting its non-linear, non-stationary, and chaotic dynamics through traditional linear methodologies [9]. The studies have shown that EEG nonlinear feature indices are significantly correlated with the power spectrum distribution in comatose and anesthetized patients [10]. The Lempel–Ziv complexity, a renowned non-linear feature, not only serves as a measure of complexity but may also act as a promising indicator of neural complexity and a biomarker for consciousness [11]. The brain, a paragon of complex systems, showcases that single-theory explanations often fall short in explaining phenomena like aging, brain electricity, and consciousness within medical research [12]. There has been an optimistic view that diseases could be targeted through specific mechanisms. However, as per complex systems theory, the cells and molecules demonstrate complex and non-linear features during disease progression [13]. This inherent unpredictability of diseases underscores the importance of understanding the variability in the effectiveness and safety of interventions among different population groups. Therefore, investigating the influence of EEG's non-linear feature indices on POD through the lens of complex systems theory promises to enrich our comprehension of various complex system issues, like disease pathology and brain functionalities [14].

In the realm of AI (artificial intelligence), ML (machine learning) and its subset, DL (deep learning), stand out as key methodologies [15]. The evolution of computer algorithms has recently amplified the application of CNN (convolutional neural networks) in deep learning. CNNs are particularly adept at extracting features from complex data types, such as images and videos, allowing for the effective utilization of patient radiological data [16, 17]. This study proposes leveraging intraoperative EEG non-linear feature indices for POD prediction, aiming to achieve precise stratification of patients at high risk for POD. This approach not only contributes to the specific field of clinical anesthesia but also paves the way for broader applications of DL in clinical research.

Methods

Patients' recruitment

This research was conducted as a prospective, observational study, adhering to the STROBE reporting guidelines [18]. The study received registration at chictr.org (ChiCTR2400080835) and approval from the Ethics Committee of the Second Affiliated Hospital of Nanjing Medical University (2024-KY-011–01). Prior to recruitment, informed consent was obtained from all participants, either directly or through a legally authorized representative.

The inclusion criteria were: (1) Age ≥ 65 years; (2) ASA (american society of anesthesiologists) classification levels I to IV; (3) Expected hospital stay of at least 3 days; (4) Scheduled for elective orthopedic surgery; (5) Expected duration of the anesthesia operation is 90 min or longer. The exclusion criteria were as follows: (1) Inability to communicate due to coma, severe dementia, or language disorders; (2) Severe cardiac, cerebral, renal, or hepatic functional abnormalities; (3) History of major trauma or major surgery within the past year; (4) History of significant organ dysfunction; (5) Preexisting conditions such as schizophrenia, epilepsy, Parkinson's disease, or myasthenia gravis.

Evaluation of POD

POD involves fluctuations in mental status, necessitating multiple repeated evaluations [19]. Trained investigators assessed patients with the 3D-confusion assessment method (3D-CAM) or the CAM for the Intensive Care Unit (CAM-ICU) twice daily for the first three postoperative days, including those unable to speak in the intensive care unit (e.g., patients still intubated) [20, 21]. This assessment encompasses four key components: (1) alterations in the level of consciousness, (2) acute fluctuations in mental status, (3) disordered thinking, and (4) inattention. POD was identified based on the presence of both (1) and (2), along with either (3), (4), or both.

Anesthetic methods

In this study, anesthesia practitioners were not informed of the details of patients' EEG waveforms and processed EEG (pEEG) data. Consequently, neither pEEG nor EEG waveforms were utilized to adjust the depth of anesthesia. Heart rate and blood pressure were maintained within 20% of baseline values, employing vasoactive and cardioactive drugs as per the anesthesiologist's discretion. To reduce potential confounding effects on EEG spectrum assessment due to EEG suppression, the use of alternative non-propofol intravenous hypnotic agents (such as midazolam, ketamine, or dexmedetomidine) was purposely avoided.

Utilization of EEG monitoring and data analysis

The monitoring of patient brain function was conducted using the ConView® system, developed by Pearlcare Medical Technology Company Limited, based in Zhejiang, China.

This system utilizes an array of electrodes placed on the forehead, specifically at positions Fp1 or Fp2. An earth electrode was positioned at Fpz, while a reference electrode was situated in the temple area, between the eye's corner and the hairline. The electrodes were calibrated to maintain a skin contact impedance below 5 kΩ for each channel, with a sampling rate set at 500 Hz. Analysis of the EEG data was carried out using MATLAB (version 2021a). The raw EEG data underwent bandpass filtering in the range of 0.5 to 50 Hz. Bad signals from the filtered data were manually removed (Such as short time amplitude reduction, signal noise and signal gaps). Firstly, the entire intraoperative EEG signal was analyzed as a whole. The entire intraoperative EEG signals were then split into 20 epochs for analysis. We calculated 14 nonlinear feature indices using a custom MATLAB script. Finally, the different complexity indices are integrated into heatmaps for DL modeling (Fig. 1).

Calculation of non-linear EEG parameters

In this study, 14 non-linear features were extracted from the intraoperative EEG data to explore their association with POD. Among these, PLZC (permutation Lempel–Ziv complexity) was calculated with an arrange dimension of 4 and an embedded dimension of 24. This measure captures the complexity of the signal by quantifying the number of distinct patterns in consecutive values, with lower PLZC values indicating reduced complexity. WE (wavelet entropy) was computed using the Daubechies 4 (db4) wavelet basis function with four decomposition layers, providing a measure of randomness and disorder across multiple time scales. Similarly, HilEn (hilbert-yellow spectral entropy) was calculated within the frequency band of 1–30 Hz, with a frequency resolution of 0.5 Hz, leveraging the instantaneous frequency spectrum to assess time-varying signal complexity.

Spectral power analysis was conducted using H transform parameters to estimate the Hurst exponent. The frequency spectrum was divided into 20 bands, starting at 1 Hz, with each band spanning 2.5 Hz. The PSD (power spectral density) across these bands was analyzed to determine the relationship between frequency and power. From this, the H (hurst exponent) was calculated, providing a measure of the long-range correlations and self-similarity in the EEG signal. AE (approximate entropy) and SampEn (sample

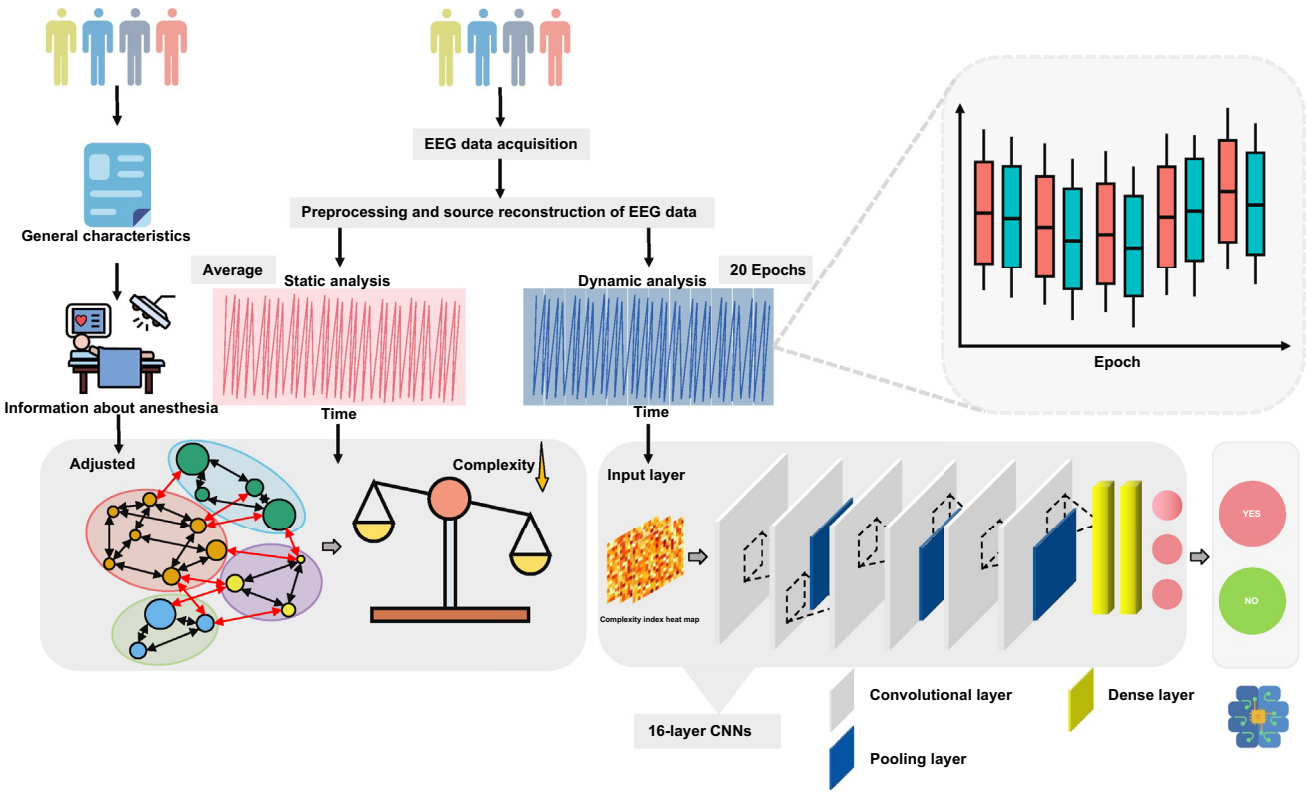

Fig. 1 Workflow of EEG non-linear feature analysis and prediction model development. The intraoperative EEG data were analyzed through static and dynamic approaches. Static analysis calculated a single non-linear feature value for each patient over the entire intraoperative period, adjusted for related factors. Dynamic analysis divided the period into 20 epochs, with non-linear features from each epoch used in a 16-layer convolutional neural network to predict outcomes, evaluated by receiver operating characteristic curves. CNN, convolutional neural network; Epoch refers to the division of the entire intraoperative period into 20 time segments

entropy) were both calculated using an embedded dimension of 2 and a similarity tolerance of 0.15 times the standard deviation of the signal. AE measures the predictability of the signal, where lower values indicate more regular patterns, while SampEn is a refined version of AE that excludes self-matching. PE (permutation entropy), another non-linear metric, was computed using an embedding dimension of 3 and a time delay of 1. PE quantifies the complexity of the signal based on the order relationships between consecutive data points.

Statistical analyses

In this study, statistical analyses were carried out using R software, version 4.2.1. For data conforming to normal distribution, the mean ± standard deviation was used for representation, and comparisons between such groups were conducted using the two independent samples t-test. Non-normally distributed data were presented as median and interquartile ranges, with the Mann–Whitney U test employed for intergroup comparisons. Count data were denoted as numbers (percentages) and compared using the chi-square test or Fisher's exact test. Linear mixed-effects model was employed to compare repeated measures data (14 nonlinear feature indices). The analysis was conducted using the lmerTest package within the R software environment. In this model, group and time (modeled as categorical variables), as well as the interaction between group and time, were treated as fixed effects. Random effects were incorporated as random intercepts for subjects. Multivariate logistic regression analysis was carried out with variables with p-value less than 0.05.

To extract the intraoperative EEG nonlinear feature indices, we used custom MATLAB scripts and routines adapted from the EEGLab toolbox. The study leveraged the VGG16 architecture within PyTorch, incorporating 13 convolutional layers, 5 pooling layers, and 4 fully connected layers. The batch size for the model analysis was set to 1, meaning one image was processed at a time. The training phase involved classifying input as 'yes' or 'no' for the occurrence of POD. Around 70% of images were utilized for training, while the remaining 30% were allocated to the validation set for CNN processing. Prior to validation, all images were standardized and resized to 224 × 224 pixels, with 3 color channels for analysis in the CNN.

Hyperparameter settings in the CNN were as follows: convolutional layers with a 3 × 3 kernel size, a stride of 1, and the ReLU activation function; pooling layers with a 2 × 2 kernel size, a stride of 2, and also using the ReLU activation function. These parameters were integrated into the softmax function to set the final fully connected layer size to 2. The CNN's learning rate was established at 0.001 to reduce the risk of overfitting and model-related errors. The model underwent a rigorous process of 1000 training iterations. Post-training, the CNN model's internal validation was conducted using the designated validation images. Finally, the predictive value of the EEG nonlinear feature indices for POD was determined by plotting the area under the receiver operating characteristic curve (AUROC).

Results

Baseline characteristics

Out of the 302 patients initially considered for the study, thirteen were excluded: three due to hearing impairments that hindered communication, three were unable to provide informed consent, four due to missing EEG data, and three whose surgeries were cancelled. Ultimately, 289 patients were included in the final analysis.

In this analysis of 289 patients, 34 individuals experienced POD, accounting for 11.8%. The basic characteristics of these patients are detailed in Table 1. No significant differences were observed in terms of educational level, preoperative pain scores, and anesthesia duration. Compared to non-POD patients, those with POD exhibited higher age, ACCI (age-adjusted charlson comorbidity index), frailty scores, ASA classifications, and ICU admission rates, while presenting lower BMI (body mass index) and MMSE (minimental state examination) scores and PLZC. Additionally, the patients undergoing femoral or hip joint surgeries were more susceptible to developing POD ($p < 0.001$).

PLZC prediction for POD

Figure 2A illustrates the changes in PLZC across groups over various epochs. Patients with POD exhibited lower PLZC levels across several epochs. During the various epochs, changes in the other 13 variables representing the nonlinear feature index between the two groups are detailed in the Supplementary Material. Upon calculating the mean PLZC values across different epochs, the PLZC levels in the POD group were found to be lower than those in the non-POD group (Fig. 2B, p = 0.031). ROC (receiver operating characteristic curve) analysis for PLZC's ability to predict POD demonstrated an AUC (area under the curve) of 0.615 (95% CI 0.517–0.711, Fig. 2C). Furthermore, after controlling for confounding factors, PLZC was identified as an independent risk factor for POD [OR = 0.210; 95% CI (0.050–0.850); $p = 0.029$] (Fig. 3).

Nonlinear features prediction for POD

To delve deeper into the predictive impact of the nonlinear feature indices on POD across various epochs, we

Table 1 General characteristics of patients between the POD and non-POD groups

Variables	Total ($n=289$)	Non-POD ($n=255$)	POD ($n=34$)	p
Gender, n (%)				0.509
Female	185 (64)	161 (63)	24 (71)	
Male	104 (36)	94 (37)	10 (29)	
Age, median (Q1,Q3)	72 (68, 77)	71 (68, 76)	80 (75, 84.75)	< 0.001
Education, median (Q1,Q3)				0.195
Elementary grade	168 (58)	144 (56)	24 (71)	
Medium grade	99 (34)	92 (36)	7 (21)	
High grade	22 (8)	19 (7)	3 (9)	
BMI, mean ± SD	24.3 ± 3.88	24.51 ± 3.87	22.77 ± 3.66	0.014
Preoperative NRS, median (Q1,Q3)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0.742
Smoking history, n (%)				0.687
No	226 (78)	198 (78)	28 (82)	
Yes	63 (22)	57 (22)	6 (18)	
Drinking history, n (%)				0.100
No	228 (79)	197 (77)	31 (91)	
Yes	61 (21)	58 (23)	3 (9)	
Coronary heart disease, n (%)				1.000
No	250 (87)	220 (86)	30 (88)	
Yes	39 (13)	35 (14)	4 (12)	
Hypertension, n (%)				0.814
No	103 (36)	92 (36)	11 (32)	
Yes	186 (64)	163 (64)	23 (68)	
Diabetes, n (%)				0.418
No	224 (78)	200 (78)	24 (71)	
Yes	65 (22)	55 (22)	10 (29)	
MMSE, median (Q1,Q3)	28 (25, 29)	28 (25.5, 29)	26 (22.25, 27.75)	< 0.001
ACCI, median (Q1,Q3)	4 (3, 5)	4 (3, 5)	4 (3.25, 5)	0.008
FRAIL scale, median (Q1,Q3)	2 (0, 3)	2 (0, 3)	2 (1.25, 3.75)	< 0.001
Surgery type, n (%)				< 0.001
Hip replacement	49 (17)	42 (16.5)	7 (20.6)	
Knee replacement	27 (9.3)	26 (10.2)	1 (2.9)	
Femoral fracture	36 (12.5)	21 (8.2)	15 (44.1)	
Lumbar fracture	177 (61.2)	166 (65.1)	11 (32.4)	
ASA, median (Q1,Q3)	2 (2, 3)	2 (2, 2.5)	3 (2, 3)	< 0.001
Anesthesia duration, median (Q1,Q3)	160 (135, 195)	165 (135, 195)	145 (127.75, 204.5)	0.437
ICU admission, n (%)				< 0.001
No	265 (92)	242 (95)	23 (68)	
Yes	24 (8)	13 (5)	11 (32)	
Length of stay, median (Q1,Q3)	10 (8, 12)	10 (8, 12)	11 (9, 13)	0.147
PLZC, median (Q1,Q3)	0.64 (0.63, 0.64)	0.64 (0.63, 0.64)	0.63 (0.62, 0.64)	0.031
B, median (Q1,Q3)	2.21 (2.07, 2.31)	2.21 (2.07, 2.31)	2.19 (2.07, 2.29)	0.638
$LZC_{k\text{-}means}$, median (Q1,Q3)	0.92 (0.90, 0.93)	0.92 (0.90, 0.93)	0.91 (0.90, 0.92)	0.190
LZC_{mean}, median (Q1,Q3)	0.92 (0.90, 0.93)	0.92 (0.90, 0.93)	0.91 (0.90, 0.92)	0.225
LZC_{median}, median (Q1,Q3)	0.90 (0.88, 0.91)	0.90 (0.88, 0.91)	0.90 (0.88, 0.91)	0.267
$LZC_{mid\text{-}point}$, median (Q1,Q3)	**0.75** (**0.74**, **0.77**)	**0.75** (**0.74**, **0.77**)	**0.76** (**0.74**, **0.77**)	0.893
A, median (Q1,Q3)	40.56 (26.58, 55.26)	40.56 (26.37, 55.07)	39.85 (27.19, 55.45)	0.764
AE, median (Q1,Q3)	1.15 (1.14, 1.15)	1.15 (1.14, 1.16)	1.15 (1.14, 1.15)	0.430
H, median (Q1,Q3)	**0.60** (**0.54**, **0.65**)	**0.61** (**0.54**, **0.65**)	**0.60** (**0.54**, **0.65**)	0.638
HilEn, median (Q1,Q3)	5.21 (5.17, 5.26)	5.21 (5.17, 5.26)	5.21 (5.18, 5.26)	0.961
PE, median (Q1,Q3)	0.92 (0.92, 0.92)	0.92 (0.92, 0.92)	0.92 (0.91, 0.92)	0.083

Table 1 (continued)

Variables	Total ($n=289$)	Non-POD ($n=255$)	POD ($n=34$)	p
r, median (Q1,Q3)	0.49 (0.25, 0.60)	0.50 (0.25, 0.61)	0.47 (0.27, 0.58)	0.751
SampEn, median (Q1,Q3)	1.17 (1.14, 1.22)	1.17 (1.14, 1.22)	1.16 (1.14, 1.20)	0.189
WE, median (Q1,Q3)	1.38 (1.33, 1.41)	1.39 (1.33, 1.42)	1.37 (1.33, 1.41)	0.553

BMI body mass index; *MMSE* mini-mental state examination; *NRS* numeric rating scale; *ACCI* age-adjusted charlson comorbidity index; *PLZC* permutation Lempel–Ziv complexity; *LZC* Lempel–Ziv complexity; *PE* permutation entropy; *SampEn* sample entropy; *AE* approximate entropy; *H* hurst exponent, ($H=(B\text{-}1)/2$); *WE* wavelet entropy; *HilEn* hilbert-yellow spectral entropy; *r* fit correlation coefficient; *A* power-law scaling coefficient; *B* power-law exponent; *POD* postoperative delirium; *ASA* american society of anesthesiologists; *FRAIL* fatigue, resistance, ambulation, illness, and loss of weight scale; Notes: $LZC_{k\text{-}means}$: The signal is divided into two clusters using k-means clustering. A binary sequence is generated by assigning values based on the proximity to the centroids. The *LZC* is then computed for the resulting binary sequence, representing the complexity of the signal. LZC_{mean}: The signal values are compared with the mean amplitude of the signal. A binary sequence is generated and its *LZC* is computed based on the number of unique subsequences. LZC_{mean}: The signal values are compared with the median amplitude. A binary sequence is created, and the *LZC* is calculated for the binary sequence. $LZC_{mid\text{-}point}$: The signal values are compared with the midpoint between the minimum and maximum values of the signal. A binary sequence is generated, and the *LZC* is computed based on the sequence's complexity

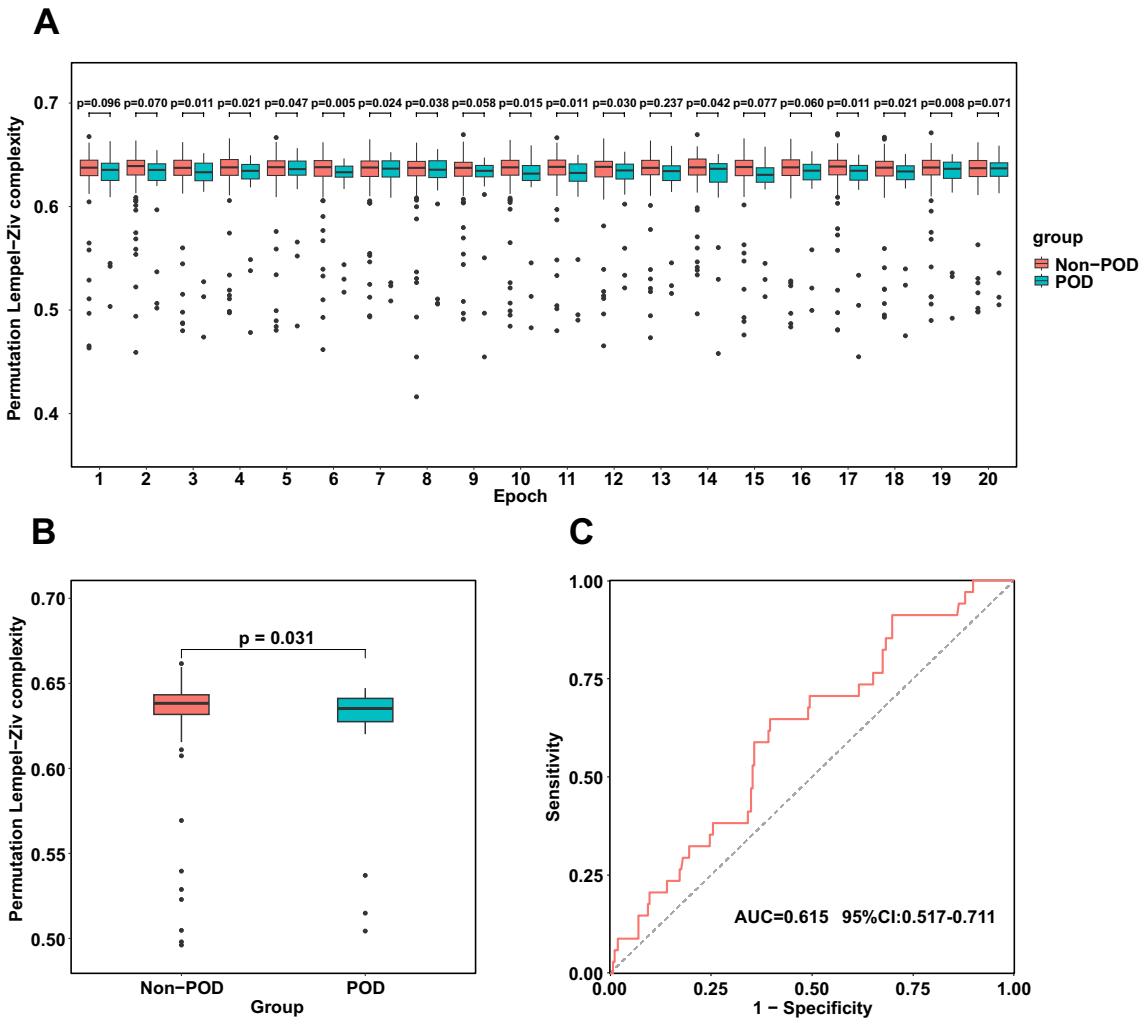

Fig. 2 Distribution of the permutation Lempel–Ziv complexity profiles. **A** Comparative analysis of permutation Lempel–Ziv complexity in POD and Non-POD groups across various epochs. **B** Comparative analysis of permutation Lempel–Ziv complexity in POD and Non-POD groups. **C** ROC curve analysis for assessing permutation Lempel–Ziv complexity in the POD group

randomly allocated 289 patients into training and validation sets at a 7:3 ratio. The two variables, coronary heart disease and PE, showed significantly different between the two groups ($p < 0.05$). Details of the training and validation sets are provided in Table 2. Upon incorporating the nonlinear feature indices across various epochs, the predictive efficacy of the nonlinear feature indices for POD was enhanced [AUC = 0.757; 95% CI 0.644–0.871] (Fig. 4A). In the validation set, the ROC curve for the nonlinear feature index's prediction of POD was 0.701 (95% CI 0.541–0.862, Fig. 4B).

Discussion

In this study, we observed a lower incidence of POD in elderly orthopedic surgery patients, which somewhat contradicts previous findings [22, 23]. This may be due to continuous advancements in anesthesia techniques and the study center's preference for combined spinal-epidural anesthesia in elderly patients with poor physical status. Consequently, anesthesiologists may opt for general anesthesia in cases with milder conditions. This may explain the lower incidence of POD observed in our study. In a post-hoc analysis, it was found that elderly orthopedic surgery patients with an ASA score of 3 or higher comprised 29% of the total study sample, aligning with our speculation.

In this prospective cohort study, our findings are as follows: (1) After adjusting for confounding factors, a decrease in intraoperative PLZC was independently associated with an increased risk of POD; (2) Taking temporal factors into account, the EEG complexity indices moderately predict POD. Collectively, these insights suggest that intraoperative EEG nonlinear feature indices should be considered a crucial factor in intraoperative monitoring, particularly for elderly patients. EEG, which records the electrical activity of brain neurons to monitor the brain's functional state, is a medical examination method for detecting electrical activity on the scalp surface [24]. Clinically, it is advised to routinely conduct EEG monitoring during surgery to reduce the risk of POD. PLZC, a nonlinear dynamic indicator, detects the likelihood of new patterns emerging in a time series, identifying variations in EEG signals without recognizing amplitude levels or frequency sizes [25]. PLZC is a reliable indicator of the complexity of brain frequency band oscillation signals, extensively used in detecting Alzheimer's disease, schizophrenia, mood disorders, and alterations in consciousness [26–29]. Although our study identifies low PLZC as a potential predictor for POD, the question remains as to how this information can be translated into actionable strategies to prevent POD when a patient is predicted to progress to this condition. While no definitive interventions are currently available, several potential approaches may warrant consideration. First, optimization of intraoperative anesthetic depth is critical. Excessive anesthetic depth has been associated with an increased risk of POD, as it can lead to reduced brain activity and impaired neuronal function [30]. Monitoring and maintaining an appropriate anesthetic depth through EEG-guided strategies, particularly those informed by nonlinear features such as PLZC, may help to minimize this risk. Second, avoiding excessive sedation and unnecessary exposure to deliriogenic medications (e.g., benzodiazepines and anticholinergic drugs) could be another strategy. Tailoring medication regimens to each patient's specific needs may reduce the likelihood of POD, especially in high-risk individuals identified through EEG-based predictions. Further studies are needed to investigate whether individualized anesthetic titration guided by EEG complexity measures, such as PLZC, can directly reduce POD incidence. Developing algorithms or clinical pathways based on EEG findings may provide a promising avenue for integrating predictive analytics with intraoperative decision-making.

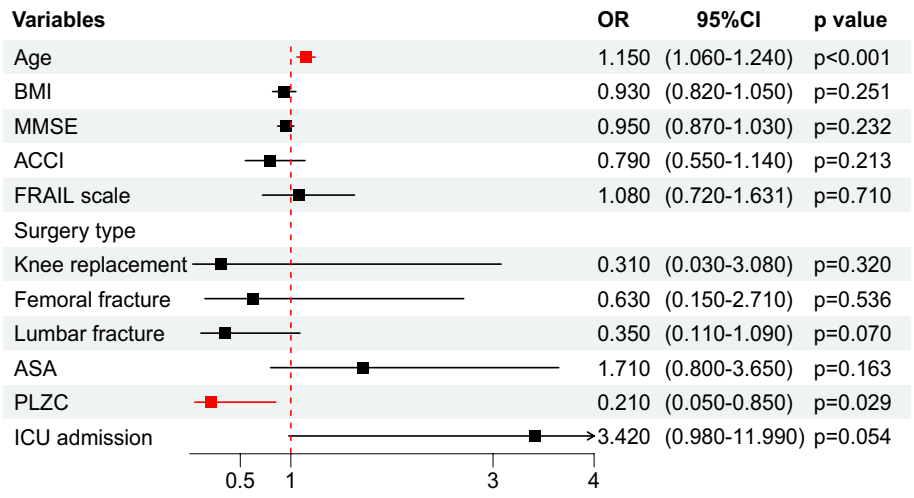

Fig. 3 Multivariate logistic regression analysis. BMI, body mass index; MMSE, mini-mental state examination; ACCI, age-adjusted charlson comorbidity index; ASA, american society of anesthesiologists; PLZC, permutation Lempel–Ziv complexity

Table 2 Characteristics of patients in training and validation data set

Variables	Total ($n=289$)	Training set ($n=201$)	Validation set ($n=88$)	p
Gender, *n* (%)				0.964
Female	185 (64)	128 (64)	57 (65)	
Male	104 (36)	73 (36)	31 (35)	
Age, median (Q1,Q3)	72 (68, 77)	71 (68, 76)	73 (68, 79.25)	0.120
Education, *n* (%)				0.855
Elementary grade	168 (58)	119 (59)	49 (56)	
Medium grade	99 (34)	67 (33)	32 (36)	
High grade	22 (8)	15 (7)	7 (8)	
BMI, mean ± SD	24.3 ± 3.88	24.36 ± 3.55	24.16 ± 4.58	0.710
Preoperative NRS, median (Q1,Q3)	**0 (0, 0)**	**0 (0, 0)**	**0 (0, 1)**	**0.150**
Smoking, *n* (%)				0.684
No	226 (78)	159 (79)	67 (76)	
Yes	63 (22)	42 (21)	21 (24)	
Drinking, *n* (%)				0.981
No	228 (79)	158 (79)	70 (80)	
Yes	61 (21)	43 (21)	18 (20)	
Coronary heart disease, *n* (%)				0.044
No	250 (87)	168 (84)	82 (93)	
Yes	39 (13)	33 (16)	6 (7)	
Hypertension, *n* (%)				0.270
No	103 (36)	67 (33)	36 (41)	
Yes	186 (64)	134 (67)	52 (59)	
Diabetes, *n* (%)				0.692
No	224 (78)	154 (77)	70 (80)	
Yes	65 (22)	47 (23)	18 (20)	
MMSE, median (Q1,Q3)	28 (25, 29)	28 (26, 29)	27 (25, 29)	0.214
ACCI, median (Q1,Q3)	4 (3, 5)	4 (3, 5)	4 (3, 5)	0.974
FRAIL scale, median (Q1,Q3)	**2 (0, 3)**	**2 (0, 3)**	**2 (1, 3)**	**0.722**
Surgery type, n (%)				0.469
Hip replacement	49 (17)	34 (17)	15 (17)	
Knee replacement	27 (9)	16 (8)	11 (12)	
Femoral fracture	36 (12)	23 (11)	13 (15)	
Lumbar fracture	177 (61)	128 (64)	49 (56)	
ASA, median (Q1,Q3)	**2 (2, 3)**	**2 (2, 3)**	**2 (2, 3)**	**0.402**
Anesthesia duration, median (Q1,Q3)	160 (135, 195)	160 (135, 195)	160.5 (133.5, 195)	0.954
ICU admission, *n* (%)				0.581
No	265 (92)	186 (93)	79 (90)	
Yes	24 (8)	15 (7)	9 (10)	
POD, *n* (%)				0.953
No	255 (88)	178 (89)	77 (88)	
Yes	34 (12)	23 (11)	11 (12)	
Length of stay, median (Q1,Q3)	10 (8, 12)	10 (8, 12)	9.5 (8, 12)	0.726
B, median (Q1,Q3)	2.21 (2.07, 2.31)	2.21 (2.07, 2.3)	2.22 (2.07, 2.31)	0.879
$LZC_{k\text{-}means}$, median (Q1,Q3)	0.92 (0.90, 0.93)	0.92 (0.90, 0.93)	0.92 (0.90, 0.93)	0.466
LZC_{mean}, median (Q1,Q3)	0.92 (0.90, 0.93)	0.92 (0.90, 0.93)	0.92 (0.90, 0.93)	0.788
LZC_{median}, median (Q1,Q3)	0.90 (0.88, 0.91)	0.90 (0.88, 0.91)	0.91 (0.89, 0.92)	0.348
$LZC_{mid\text{-}point}$, median (Q1,Q3)	**0.75 (0.74, 0.77)**	**0.75 (0.74, 0.77)**	**0.75 (0.73, 0.77)**	0.462
A, median (Q1,Q3)	40.56 (26.58, 55.26)	40.56 (26.58, 55.26)	40.47 (26.88, 54.53)	0.872
AE, median (Q1,Q3)	1.15 (1.14, 1.15)	1.15 (1.14, 1.15)	1.15 (1.14, 1.15)	0.117
H, median (Q1,Q3)	0.60 (0.54, 0.65)	0.6 (0.54, 0.65)	0.61 (0.53, 0.65)	0.879

Table 2 (continued)

Variables	Total ($n=289$)	Training set ($n=201$)	Validation set ($n=88$)	p
HilEn, median (Q1,Q3)	5.21 (5.17, 5.26)	5.21 (5.17, 5.26)	5.20 (5.17, 5.26)	0.831
PE, median (Q1,Q3)	0.92 (0.92, 0.92)	0.92 (0.92, 0.92)	0.92 (0.92, 0.92)	0.036
r, median (Q1,Q3)	0.49 (0.25, 0.6)	0.49 (0.25, 0.6)	0.50 (0.25, 0.61)	0.806
SampEn, median (Q1,Q3)	1.17 (1.14, 1.22)	1.17 (1.14, 1.22)	1.18 (1.15, 1.23)	0.060
WE, median (Q1,Q3)	1.38 (1.33, 1.41)	1.38 (1.33, 1.41)	1.39 (1.33, 1.42)	0.927
PLZC, median (Q1,Q3)	0.64 (0.63, 0.64)	0.64 (0.63, 0.64)	0.64 (0.63, 0.64)	0.434

BMI body mass index; *MMSE* mini-mental state examination; *NRS* numeric rating scale; *ACCI* age-adjusted charlson comorbidity index; *PLZC* permutation Lempel–Ziv complexity; *LZC* Lempel–Ziv complexity; *PE* permutation entropy; *SampEn* sample entropy; *AE* approximate entropy; *H* hurst exponent, ($H=(B\text{-}1)/2$); *WE* wavelet entropy; *HilEn* hilbert-yellow spectral entropy; *r*, fit correlation coefficient; *A* power-law scaling coefficient; *B* power-law exponent; *POD* postoperative delirium; *ASA* american society of anesthesiologists; *FRAIL* fatigue, resistance, ambulation, illness, and loss of weight scale; Notes: $LZC_{k\text{-}means}$: The signal is divided into two clusters using k-means clustering. A binary sequence is generated by assigning values based on the proximity to the centroids. The *LZC* is then computed for the resulting binary sequence, representing the complexity of the signal. LZC_{mean}: The signal values are compared with the mean amplitude of the signal. A binary sequence is generated and its *LZC* is computed based on the number of unique subsequences. LZC_{median}: The signal values are compared with the median amplitude. A binary sequence is created, and the *LZC* is calculated for the binary sequence. $LZC_{mid\text{-}point}$: The signal values are compared with the midpoint between the minimum and maximum values of the signal. A binary sequence is generated, and the *LZC* is computed based on the sequence's complexity

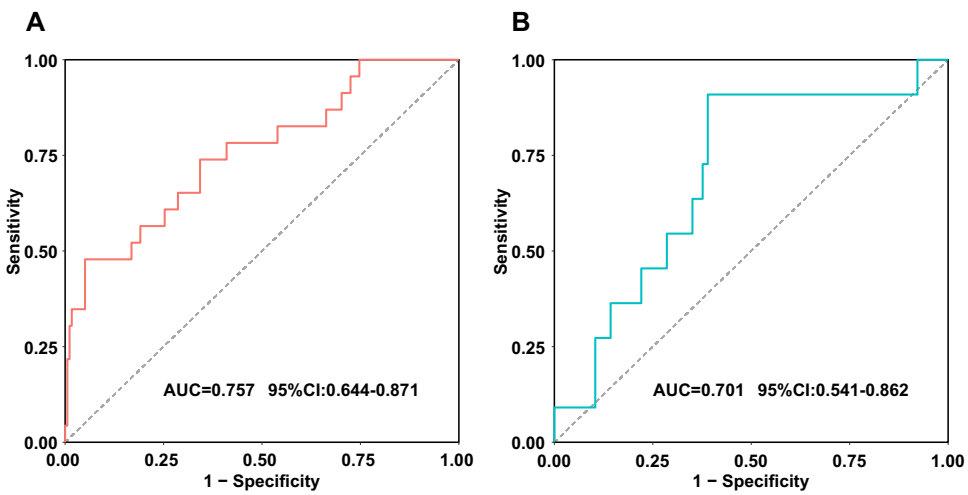

Fig. 4 Convolutional neural network-based analysis of EEG nonlinear feature index. **A** ROC curve analysis of the EEG nonlinear feature index in the training dataset. **B** ROC curve analysis of the EEG nonlinear feature index in the validation dataset

Our findings suggest that low PLZC, a non-linear EEG feature, is an independent risk factor for POD. This complements previous research by Kinoshita et al. which demonstrated that intraoperative alpha-power, a linear EEG parameter, can also predict POD development [31]. Both approaches offer valuable insights into the neurophysiological changes associated with POD, but they reflect different aspects of brain dynamics and may have distinct implications for clinical use. Intraoperative alpha-power primarily reflects cortical synchronization and is known to decrease with deeper levels of anesthesia and impaired brain function [32]. Higher alpha power during surgery has been associated with better brain health and lower POD risk [33]. The advantage of using alpha-power as a predictor lies in its simplicity and interpretability, as it represents a well-established linear EEG metric that can be easily monitored with existing intraoperative EEG systems.

In contrast, PLZC is a non-linear measure that captures the complexity and randomness of brain dynamics, providing a broader perspective on neural information processing [34]. Unlike alpha-power, PLZC is sensitive to subtle, non-linear interactions in the EEG signal that may be overlooked by linear metrics. A decrease in PLZC suggests reduced neural complexity, which could reflect impaired communication and adaptability within brain networks under anesthesia, thereby increasing the vulnerability to POD. The computation of PLZC, however, is more complex and requires advanced signal processing techniques, which may limit its immediate applicability in routine clinical settings. When comparing these two metrics, it is important to recognize

that they capture different dimensions of brain function. Alpha-power provides a direct measure of spectral activity within a specific frequency band [35], while PLZC analyzes the temporal structure and complexity of the entire EEG signal. While alpha-power may be more practical for real-time monitoring, PLZC could offer additional predictive value by uncovering hidden neural dynamics associated with POD. This distinction suggests that combining linear and non-linear features might yield a more comprehensive model for predicting POD, leveraging the strengths of both approaches.

Future research should investigate the integration of linear metrics such as alpha-power, with non-linear features like PLZC in predictive algorithms. This combined approach could improve the sensitivity and specificity of POD prediction, ultimately guiding intraoperative interventions to optimize brain health and reduce the risk of POD. Additionally, comparative studies are needed to evaluate the relative performance of these metrics across diverse patient populations and surgical settings.

Limitations

Although the findings of this study are promising, revealing lower EEG signal complexity in elderly orthopedic surgery patients as potential biomarkers for POD, some limitations persist. First, this single-center prospective study, focused solely on orthopedic surgery with a small sample size, has limited generalizability, necessitating large, independent, multicenter clinical trials to confirm the precise role of EEG nonlinear feature indices in patients with POD. Second, the study design was neither randomized nor propensity-matched, potentially allowing unmeasured confounding factors to influence our results [36, 37]. Third, patients received various drugs affecting brain neuronal activity during anesthesia, and we were unable to control for the potential confounding effects of these medications. Finally, the exclusion of certain laboratory findings may introduce bias into the results.

Conclusions

This study indicates that the EEG nonlinear feature indices could serve as a biomarker for POD and may be utilized to develop a diagnostic model for predicting the occurrence of POD in elderly patients undergoing orthopedic surgery.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00540-025-03471-4.

Acknowledgements We thanks all participants in the study.

Author contributions Jie Sun and Mu-huo Ji designed the study. Xiao-yi Hu and Yu-chen Dai collected the data and drafted the manuscript. Lan-yue Zhu analyzed the data and provided critical revision of the manuscript for important intellectual content. Jian-jun Yang critically reviewed data and participated writing manuscript. All authors have read and approved the manuscript.

Funding The trial was funded by the National Natural Science Foundation of China fund (82172131).

Data availability The datasets used and analysed in this study may be obtained from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

1. Oh ES, Fong TG, Hshieh TT, Inouye SK. Delirium in older persons: advances in diagnosis and treatment. JAMA. 2017;318(12):1161–74.
2. Evered L, Atkins K, Silbert B, Scott DA. Acute peri-operative neurocognitive disorders: a narrative review. Anaesthesia. 2022;77(Suppl 1):34–42.
3. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. Lancet. 2014;383(9920):911–22.
4. Bokeriia LA, Golukhova EZ, Polunina AG. Postoperative delirium in cardiac operations: microembolic load is an important factor. Ann Thorac Surg. 2009;88(1):349–51.
5. Guan HL, Liu H, Hu XY, Abdul M, Dai MS, Gao X, Chen XF, Zhou Y, Sun X, Zhou J, Li X, Zhao Q, Zhang QQ, Wang J, Han Y, Cao JL. Urinary albumin creatinine ratio associated with postoperative delirium in elderly patients undergoing elective non-cardiac surgery: a prospective observational study. CNS Neurosci Ther. 2022;28(4):521–30.
6. Xing H, Zhou W, Fan Y, Wen T, Wang X, Chang G. Development and validation of a postoperative delirium prediction model for patients admitted to an intensive care unit in China: a prospective study. BMJ Open. 2019;9(11): e030733.
7. Ling YT, Guo QQ, Wang SM, Zhang LN, Chen JH, Liu Y, Xuan RH, Qu B, Liu LG, Wen ZS, Xu JK, Jiang LL, Xian WB, Wu B, Zhang CM, Chen L, Liu JL, Jiang N. Nomogram for prediction of postoperative delirium after deep brain stimulation of subthalamic nucleus in Parkinson's disease under general anesthesia. Parkinsons Dis. 2022;2022:6915627.
8. Cao KX, Ma ML, Wang CZ, Iqbal J, Si JJ, Xue YX, Yang JL. TMS-EEG: an emerging tool to study the neurophysiologic biomarkers of psychiatric disorders. Neuropharmacology. 2021;197:108574.
9. Acharya UR, Sudarshan VK, Adeli H, Santhosh J, Koh JE, Puthankatti SD, Adeli A. A novel depression diagnosis index using nonlinear features in EEG signals. Eur Neurol. 2015;74(1–2):79–83.
10. Colombo MA, Napolitani M, Boly M, Gosseries O, Casarotto S, Rosanova M, Brichant JF, Boveroux P, Rex S, Laureys S, Massimini M, Chieregato A, Sarasso S. The spectral exponent of the resting EEG indexes the presence of consciousness during unresponsiveness induced by propofol, xenon, and ketamine. Neuroimage. 2019;189:631–44.
11. Brito MA, Li D, Fields CW, Rybicki-Kler C, Dean JG, Liu T, Mashour GA, Pal D. Cortical acetylcholine levels

correlate with neurophysiologic complexity during subanesthetic ketamine and nitrous oxide exposure in rats. Anesth Analg. 2022;134(6):1126–39.
12. Andersen BM, Faust Akl C, Wheeler MA, Chiocca EA, Reardon DA, Quintana FJ. Glial and myeloid heterogeneity in the brain tumour microenvironment. Nat Rev Cancer. 2021;21(12):786–802.
13. Azami H, Daftarifard E, Humeau-Heurtier A, Fernandez A, Abasolo D, Rajji TK. Assessment and comparison of nonlinear measures in resting-state magnetoencephalograms in Alzheimer's disease and mild cognitive impairment. J Alzheimers Dis. 2023;96(3):1151–62.
14. Jiao B, Li R, Zhou H, Qing K, Liu H, Pan H, Lei Y, Fu W, Wang X, Xiao X, Liu X, Yang Q, Liao X, Zhou Y, Fang L, Dong Y, Yang Y, Jiang H, Huang S, Shen L. Neural biomarker diagnosis and prediction to mild cognitive impairment and Alzheimer's disease using EEG technology. Alzheimers Res Ther. 2023;15(1):32.
15. Crampon K, Giorkallos A, Deldossi M, Baud S, Steffenel LA. Machine-learning methods for ligand-protein molecular docking. Drug Discov Today. 2022;27(1):151–64.
16. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Moreira AL, Razavian N, Tsirigos A. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med. 2018;24:1559–67.
17. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18:500–10.
18. Braillon A, Naudet F. STROBE and pre-registration of observational studies. BMJ. 2023;380:90.
19. Fialho Silva IT, Assis Lopes P, Timotio Almeida T, Ramos SC, Caliman Fontes AT, Guimarães Silva D, Martins Soares C, Oliveira Carneiro L, Souza IFB, Ferreira Abreu F, Nascimento Silva G, de MascarenhasSouza L, Brito Pinheiro T, de Souza E, Silva FN, de Santana JP, Kelly Silva B, Almeida Souza D, Silva Macedo S, Almeida Ismael LS, Pereira de Jesus PA. Impact of Delirium and Its Motor Subtypes on Stroke Outcomes. Stroke. 2021;52(4):1322–9.
20. Liu X, Nakano M, Yamaguchi A, Bush B, Akiyoshi K, Lee JK, Koehler RC, Hogue CW, Brown CH 4th. The association of bispectral index values and metrics of cerebral perfusion during cardiopulmonary bypass. J Clin Anesth. 2021;74: 110395.
21. Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, Speroff T, Gautam S, Bernard GR, Inouye SK. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). Crit Care Med. 2001;29(7):1370–9.
22. Xu X, Hu X, Wu Y, Li Y, Zhang Y, Zhang M, Yang Q. Effects of different BP management strategies on postoperative delirium in elderly patients undergoing hip replacement: a single center randomized controlled trial. J Clin Anesth. 2020;62: 109730.
23. Tao M, Zhang S, Han Y, Li C, Wei Q, Chen D, Zhao Q, Yang J, Liu R, Fang J, Li X, Zhang H, Liu H, Cao JL. 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(1):88–96.
24. Fitzgibbon SP, DeLosAngeles D, Lewis TW, Powers DM, Whitham EM, Willoughby JO, Pope KJ. Surface Laplacian of scalp electrical signals and independent component analysis resolve EMG contamination of electroencephalogram. Int J Psychophysiol. 2015;97(3):277–84.
25. Ren Z, Yue M, Han X, Zhao Z, Wang B, Hong Y, Zhao T, Wang N, Zhao P, Hong Y, Wang Q, Zhao Y. The potential of the Lempel-Ziv complexity of the EEG in diagnosing cognitive impairment in patients with temporal lobe epilepsy. Epileptic Disord. 2023;25(3):331–42.
26. Wu T, Sun F, Guo Y, Zhai M, Yu S, Chu J, Yu C, Yang Y. Spatiotemporal dynamics of entropy in EEGS during music stimulation of Alzheimer's Disease patients with different degrees of Dementia. Entropy (Basel). 2022;24(8):1137.
27. Akar SA, Kara S, Latifoğlu F, Bilgiç V. Analysis of the complexity measures in the EEG of schizophrenia patients. Int J Neural Syst. 2016;26(2):1650008.
28. Altıntop ÇG, Latifoğlu F, Akın AK, Bayram A, Çiftçi M. Classification of depth of coma using complexity measures and nonlinear features of electroencephalogram signals. Int J Neural Syst. 2022;32(5):2250018.
29. Gaggioni G, Shumbayawonda E, Montanaro U, Ly JQM, Phillips C, Vandewalle G, Abásolo D. Time course of cortical response complexity during extended wakefulness and its differential association with vigilance in young and older individuals. Biochem Pharmacol. 2021;191: 114518.
30. Fang PP, Shang ZX, Xu J, Hu J, Zhang SC, Fan YG, Lu Y, Liu XS, Maze M. Contribution of intraoperative electroencephalogram suppression to frailty-associated postoperative delirium: mediation analysis of a prospective surgical cohort. Br J Anaesth. 2023;130(2):e263–71.
31. Kinoshita H, Saito J, Kushikata T, Oyama T, Takekawa D, Hashiba E, Sawa T, Hirota K. The perioperative frontal relative ratio of the alpha power of electroencephalography for predicting postoperative delirium after highly invasive surgery: a prospective observational study. Anesth Analg. 2023;137(6):1279–88.
32. Li Z, Wang P, Han L, Hao X, Mi W, Tong L, Liang Z. Age-dependent coupling characteristics of bilateral frontal EEG during desflurane anesthesia. Physiol Meas. 2024;45(5). https://doi.org/10.1088/1361-6579/ad46e0.
33. Khalifa C, Lenoir C, Robert A, Watremez C, Kahn D, Mastrobuoni S, Aphram G, Ivanoiu A, Bonhomme V, Mouraux A, Momeni M. Intra-operative electroencephalogram frontal alpha-band spectral analysis and postoperative delirium in cardiac surgery: a prospective cohort study. Eur J Anaesthesiol. 2023;40(10):777–87.
34. Liang Z, Shao S, Lv Z, Li D, Sleigh JW, Li X, Zhang C, He J. Constructing a consciousness meter based on the combination of non-linear measurements and genetic algorithm-based support vector machine. IEEE Trans Neural Syst Rehabil Eng. 2020;28(2):399–408.
35. Bazanova OM, Vernon D. Interpreting EEG alpha activity. Neurosci Biobehav Rev. 2014;44:94–110.
36. Evans E, Swanson MB, Mohr N, Boulos N, Vaughan-Sarrazin M, Chan PS, Girotra S. American Heart Association's Get With The Guidelines-Resuscitation investigators. Epinephrine before defibrillation in patients with shockable in-hospital cardiac arrest: propensity matched analysis. BMJ. 2021;375:066534.
37. Nevins P, Davis-Plourde K, Pereira Macedo JA, Ouyang Y, Ryan M, Tong G, Wang X, Meng C, Ortiz-Reyes L, Li F, Caille A, Taljaard M. A scoping review described diversity in methods of randomization and reporting of baseline balance in stepped-wedge cluster randomized trials. J Clin Epidemiol. 2023;157:134–45.