Zhu et al. BMC Neurology (2025) 25:373
https://doi.org/10.1186/s12883-025-04389-w

BMC Neurology

RESEARCH

Open Access

Creating a delirium risk assessment model for geriatric patients with femoral neck fractures through artificial intelligence techniques: a theoretical exploration

Yu Zhu[1†], Ning Zhou[2†], Renrui Liang[1*], Jian-Jun Yang[3*] and Cheng-Mao Zhou[1*]

Abstract

Background Postoperative delirium (POD) is a common occurrence following orthopedic surgery, particularly in the older population. However, there is a relative scarcity of research on the use of intelligent algorithms to predict POD in older patients after orthopedic surgery. Therefore, the objective of this study was to evaluate the efficacy of ten distinct intelligent algorithms in predicting POD in older patients undergoing femoral neck fracture surgery.

Methods This study selected ten advanced artificial intelligence algorithms to predict the occurrence of postoperative delirium by analyzing patient data.

Results A total of 1889 patients were included in this study. The dataset for this study was divided into a training dataset, which consisted of synthetic data, and a testing dataset, representing real-world clinical scenarios. In the training dataset, we identified 267 cases of POD, accounting for 26.70% of the group. In the testing dataset, 172 cases of POD were confirmed, representing 19.35% of the group. Analysis using the Gradient Boosting Decision Tree (GBDT) algorithm revealed that age, preoperative hemoglobin levels, duration of anesthesia, and intraoperative blood loss are key predictive factors for POD in older patients with femoral neck fractures. Among the intelligent algorithms tested for predicting POD in the testing group, logistic regression, random forest, and the Multilayer Perceptron Classifier (MLPC) performed best with accuracy rates of 0.810, 0.810, and 0.808, respectively. In terms of precision, MLPC led with a score of 1.000, followed by random forest (0.714) and logistic regression (0.548). The highest recall rates were achieved by Gaussian Naive Bayes (gnb, 0.337) and AdaBoost (adab, 0.198). Gaussian Naive Bayes also performed best in F1 score (0.244). In the evaluation of the Area Under the Curve (AUC), logistic regression, MLPC, and XGBoost (XGB) demonstrated the best performance, with values of 0.669, 0.669, and 0.652, respectively.

†Yu Zhu and Ning Zhou contributed equally to this work.

*Correspondence:
Renrui Liang
lrr011@126.com
Jian-Jun Yang
yjyangjj@126.com
Cheng-Mao Zhou
zhouchengmao187@foxmail.com

Full list of author information is available at the end of the article

Conclusions The results of this study indicate that the Multilayer Perceptron Classifier (MLPC) algorithm performed the most excellently in predicting POD after femoral neck fracture surgery in older adults, with an accuracy rate reaching 80.8%. These findings suggest that machine learning algorithms, particularly MLPC, have significant potential and practical effectiveness in predicting POD in specific older patient populations.

Keywords Machine learning, POD, MLPC, Predicting, Older

Introduction

Postoperative delirium (POD), a pervasive and acute neurocognitive disorder, typically manifests in the immediate postoperative phase, particularly among geriatric patients. Characterized by fluctuations in consciousness, attention, cognitive function, and perceptual disturbances, POD stands as one of the most prevalent and hazardous complications following surgery in older adults [1]. One Study have indicated that the incidence of POD after major orthopedic procedures can reach 16%, with particularly high-risk groups, such as elderly patients undergoing major orthopedic surgery, experiencing rates as high as 53% [2, 3]. POD is not only a critical concern in its own right but also a precursor to a cascade of secondary complications, including pressure ulcers, fall-related injuries, respiratory distress, urinary tract infections, myocardial infarctions, and atrial fibrillations, collectively contributing to increased mortality rates [4]. Furthermore, POD can precipitate delayed recovery, extended hospital stays, augmented healthcare expenditures, and post-discharge decrements in rehabilitation efficacy and quality of life [5]. The prevention and management of POD remain non-standardized, largely attributed to the enigmatic etiology and multifactorial risk factors involved. Isolating the precise high-risk factors precipitating POD is essential for elucidating its pathogenesis. Consequently, the development of robust predictive models is imperative for guiding preventative and therapeutic strategies aimed at mitigating POD incidence [6].

The early identification and intervention of POD present a formidable and multifaceted challenge in clinical practice. However, the burgeoning integration of artificial intelligence (AI) within the medical field promises to augment clinical decision-making and enhance patient outcomes through "intelligent" machine assistance. AI, predominantly operationalized through machine learning, entails the extraction of insights and acquisition of novel knowledge from complex datasets. Research has demonstrated the potential of machine learning algorithms in predicting POD [7], post-laparoscopic surgery patient ileus [8], and estimating blood loss during liver cancer surgery [9]. Despite these advances, there is a dearth of literature concerning intelligent algorithms for POD prediction specifically in the context of orthopedic surgery.

This study is explicitly designed to address this research gap by scrutinizing the efficacy of ten state-of-the-art intelligent algorithms in predicting POD among older adults who have undergone femoral neck fracture surgery. By doing so, we aim to forge a clearer understanding of the risk factors associated with POD and develop more accurate predictive models, thereby informing clinical strategies to preempt and manage this debilitating condition more effectively.

Materials and methods

Patient data characteristics

This study analyzed patient data for individuals aged 65 and over who underwent elective femoral neck fracture surgery from January 2016 to December 2020. The data analysis excluded records based on the following criteria: (1) conservative treatment; (2) bilateral surgery or non-thigh neck surgery; (3) multiple trauma or fracture; (4) lack of Hb concentration or blood transfusion data, lack of cognitive or delirium assessment; and (5) preoperative cognitive impairment. The variables included laboratory hemoglobin concentrations, demographic characteristics such as sex, age, body mass index, and history of smoking and drinking, and the American Society of Anesthesia (ASA) score. The burden of complications includes severe heart disease, lung disease, metabolic disease, and cerebrovascular disease. Data on the time of operation and anesthesia, the amount of blood lost during operation, and the type of operation and anesthesia were collected. In addition, information on perioperative drug use was collected. Delirium was assessed with (CAM-ICU) [10] and CAM [11] methods. In patients over 65 years old, mini mental state examination (MMSE) [12] was used to evaluate cognitive function regularly during preoperative visits, and mild cognitive impairment was recorded if the MMSE score was < 27. Finally, patients without mild cognitive impairment were divided into POD group and non-POD group.

Artificial intelligence algorithm

In our study, we've adopted a diverse set of contemporary machine learning and deep learning algorithms, including Logistic Regression for binary classification, Decision Tree for straightforward decision-making, Random Forest for reducing overfitting through ensemble learning, Gradient Boosting Decision Tree (GBDT) for sequential model optimization, Light Gradient Boosting Machine

(LGBM) for efficient large-scale data processing, Extreme Gradient Boosting (XGB) for handling extensive machine learning tasks, Multilayer Perceptron Classifier (MLPC) for capturing non-linear data patterns, Gaussian Naive Bayes (GNB) for probabilistic classification in high-dimensional spaces, K-Nearest Neighbors (KNN) for classifying based on feature space proximity, and AdaBoost (ADAB) for creating robust classifiers from weak learners. By integrating these sophisticated techniques, our research endeavors to offer a more precise and up-to-date application of machine learning methodologies.

This study utilized the Synthetic Data Vault(SDV) algorithm to generate virtual data based on real data. These synthetic data mimic the statistical characteristics and patterns of the original data, ensuring that the new dataset maintains the integrity and distribution features of the real data. To evaluate the quality and fidelity of the synthetic data, a comprehensive set of similarity metrics was used. These metrics cover a range from "0–1", quantifying the degree of similarity between synthetic data and actual data. In this scoring system, a score of "0" indicates no similarity, indicating data synthesis failure, while a score of "1" represents a perfect match, indicating that the synthetic data is indistinguishable from real data.

Firstly, PERSON correlation analysis was used to explore the correlation between variables. Subsequently, we utilized the virtual data as the training cohort to train and construct the model, followed by performance validation on real-world data. Moreover, it has been verified by a 5-fold cross-verification method. Manual parameter adjustment and grid parameter adjustment methods were used to optimize the model. Indicators were used to evaluate the model including C statistics (that is, the area under the receiver's working characteristics [ROC] curve), precision, accuracy, recall rate and F1 score. To understand the contribution of each variable to the machine learning model, the importance of each variable was calculated and the results were shown in graphs. All the intelligent algorithms were carried out on the Python software.

General statistical analysis

General statistical analysis was performed by R statistical software package. The measurement data in accordance with the normal distribution were shown as $\bar{x} \pm s$, and the comparison between the two groups was expressed by t test. The counting data were expressed by the number of cases (%), the comparison between two groups was conducted by χ^2 test. The test level α value was taken as bilateral 0.05. $P<0.05$ was considered statistically different.

The study's design flowchart can be found in Supplementary Fig. 1.

Results

A total of 1889 patients were included in this study. The Synthetic Data Vault (SDV)algorithm evaluated the similarity between the two datasets, the virtual dataset (training group, $N=1000$) and the real-world dataset (testing group, $N=889$), to be 91%. (Table 1) This means that the synthetically generated data can be used for modeling real data. And in the training cohort, which represents the synthetic and real data synthesis group, 267 individuals (26.70%) were confirmed to have delirium, whereas in the testing cohort, which corresponds to the real-world group, 172 individuals (19.35%) were diagnosed with delirium.

Correlation between POD variables

Correlation analysis showed that there was a positive correlation between opioid use and intraoperative bleeding and POD, and a negative correlation between preoperative hemoglobin and early POD.(Fig. 1).

The feature engineering results of the Gradient Boosting Decision Tree (GBDT) algorithm reveal that the occurrence of Postoperative Delirium (POD) in elderly patients with femoral fractures is closely associated with multiple factors. As depicted in Fig. 2, the correlation of each factor is represented by the height of the circles, with higher circles indicating a closer relationship to postoperative delirium.

Specifically, the analysis shows that age, preoperative hemoglobin levels, duration of anesthesia, and intraoperative blood loss are the main factors influencing the occurrence of postoperative delirium. In Fig. 2, the higher position of the circles for these factors indicates a significant association with the occurrence of postoperative delirium. Among them, older patients have an increased risk of postoperative delirium; lower preoperative hemoglobin levels may increase the likelihood of delirium; longer anesthesia duration may impact the patient's nervous system, thereby raising the risk of delirium; and increased intraoperative blood loss may also lead to physiological instability in patients after surgery, increasing the risk of delirium.

In the realm of predicting Postoperative Delirium (POD), intelligent algorithms have demonstrated varying degrees of effectiveness when applied to a training group. The performance of these algorithms was evaluated based on several key metrics, including accuracy, precision, recall rate, F1 score, and Area Under the Curve (AUC). The top performers in the training group were as follows (Table 2 and Fig. 3):

Accuracy

The algorithms that demonstrated the highest accuracy in predicting POD were the k-Nearest Neighbors (knn) with a score of 0.754 and the Random Forest algorithm

Table 1 Basic characteristic information

	Training(Virtual data)	Test(real-world data)
Number	1000	889
Age(years)	77.00 (73.00–83.00)	79.00 (72.00–84.00)
BMI(kg/m2)	22.70 (3.71)	22.89 (3.69)
Pre-operative hemoglobin(g/l)	119.30 (16.81)	119.69 (16.64)
Blood transfusion(IU)	2.30 (1.10–3.70)	2.00 (2.00–3.00)
Operation time(minute)	90.00 (65.75–114.00)	85.00 (65.00-110.00)
Anesthesia time(minute)	126.50 (100.00-152.00)	120.00 (100.00-145.00)
Intraoperative blood loss(ml)	207.25 (64.34)	208.13 (64.99)
Sex		
Male	411 (41.10%)	243 (27.33%)
Female	589 (58.90%)	646 (72.67%)
Smoking		
No	874 (87.40%)	784 (88.19%)
Yes	126 (12.60%)	105 (11.81%)
Stroke history		
No	916 (91.60%)	814 (91.56%)
Yes	84 (8.40%)	75 (8.44%)
Hypertension on medicine		
No	546 (54.60%)	550 (61.87%)
Yes	454 (45.40%)	339 (38.13%)
Coronary heart disease		
No	819 (81.90%)	770 (86.61%)
Yes	181 (18.10%)	119 (13.39%)
Atrial fibrillation		
No	997 (99.70%)	864 (97.19%)
Yes	3 (0.30%)	25 (2.81%)
COPD		
No	978 (97.80%)	852 (95.84%)
Yes	22 (2.20%)	37 (4.16%)
Diabetes mellitus		
No	704 (70.40%)	717 (80.65%)
Yes	296 (29.60%)	172 (19.35%)
ASA		
2	141 (14.10%)	90 (10.12%)
3	733 (73.30%)	732 (82.34%)
4	126 (12.60%)	67 (7.54%)
Hypoalbuminemia (pre)		
No	776 (77.60%)	752 (84.59%)
Yes	224 (22.40%)	137 (15.41%)
Benzodiazepines use		
No	995 (99.50%)	857 (96.40%)
Yes	5 (0.50%)	32 (3.60%)
Opioid use		
No	431 (43.10%)	319 (35.88%)
Yes	569 (56.90%)	570 (64.12%)
Midazolam use		
No	627 (62.70%)	666 (74.92%)
Yes	373 (37.30%)	223 (25.08%)
Postoperative delirium		
No	733 (73.30%)	717 (80.65%)
Yes	267 (26.70%)	172 (19.35%)

Mean(SD) Median (Q1-Q3)/*N*(%)

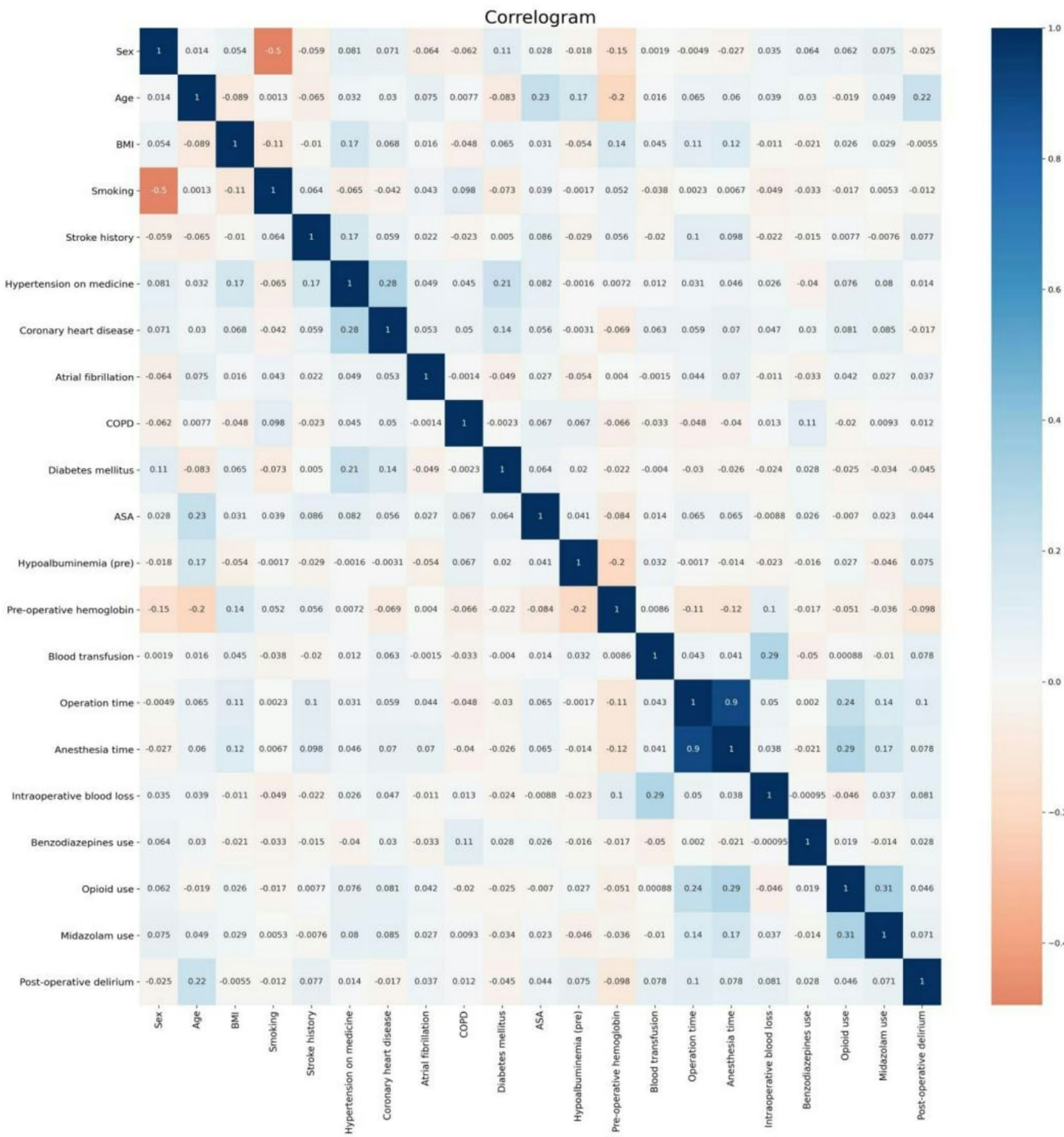

Fig. 1 Correlation between variables

with a score of 0.746. These scores indicate that these two algorithms were the most successful in correctly identifying instances of POD within the training dataset.

Precision

When it came to precision, the Extreme Gradient Boosting (XGB) and the Multilayer Perceptron Classifier (MLPC) both achieved a perfect score of 1.000, suggesting that these algorithms were highly accurate in identifying true positive cases without a significant number of false positives.

Recall rate

In terms of recall rate, which measures the ability of an algorithm to find all relevant cases, the Gaussian Naive Bayes (gnb) algorithm led with a score of 0.165, followed by knn at 0.135, and Random Forest at 0.056. These

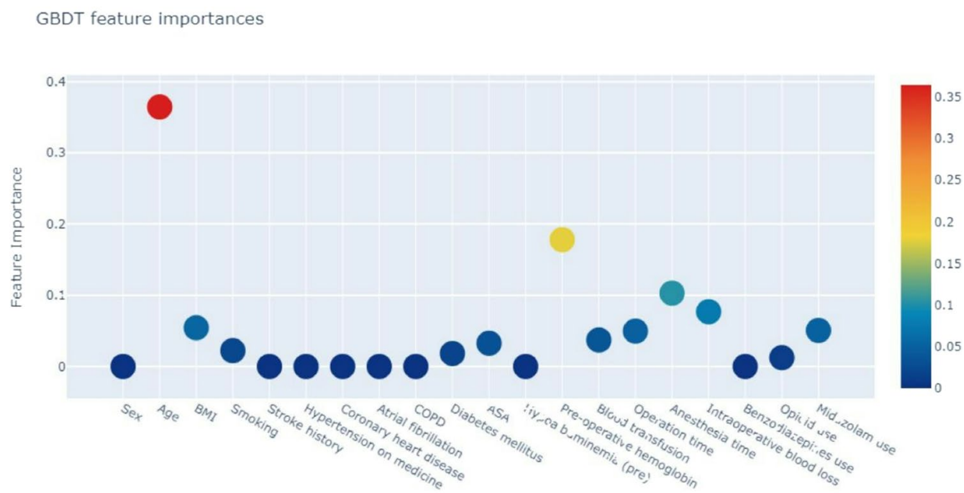

Fig. 2 Variable importance of features included in Gradient Boosting Decision Tree-GBDT algorithm for prediction of POD

Table 2 Forecast results for training group

Model name	AUC	Accuracy	Precision	Recall	F1
Logistic Regression	0.654	0.734	0.524	0.041	0.076
Decision Tree	0.618	0.733	0.000	0.000	0.000
Random Forest	0.819	0.746	0.882	0.056	0.106
Gradient Boosting	0.681	0.733	0.000	0.000	0.000
XGB	0.700	0.734	1.000	0.004	0.007
LGBM	0.682	0.733	0.000	0.000	0.000
MLPC	0.655	0.734	1.000	0.004	0.007
gnb	0.647	0.720	0.436	0.165	0.239
knn	0.742	0.754	0.706	0.135	0.226
adab	0.643	0.733	0.000	0.000	0.000

Abbreviate: Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree- Gradient Boosting, *LGBM* light gradient boosting machine, *XGB* Extreme gradient boosting, *MLPC* Multilayer Perceptron Classifier, *gnb* Gaussian naive Bayes, *knn* K-nearst neighbors, *ada* AdaBoost

results highlight the algorithms' effectiveness in capturing true positive cases.

F1 score

The F1 score, which balances precision and recall, saw gnb at the top with a score of 0.239, knn at 0.226, and Random Forest at 0.106. These scores reflect a balance between the algorithms' precision and recall capabilities.

AUC

The AUC, which measures the model's ability to discriminate between classes, had Random Forest at the forefront with a score of 0.819, knn at 0.742, and XGB at 0.700. These scores indicate the algorithms' overall predictive power in distinguishing between patients who developed POD and those who did not.

When these intelligent algorithms were applied to a test group for predicting POD, the results were as follows(Table 3 and Fig. 4):

Accuracy

The Logistic Regression, Random Forest, and MLPC algorithms were the top performers in terms of accuracy, each achieving a score of 0.810, 0.810, and 0.808, respectively. This indicates that these algorithms were highly effective in correctly predicting POD in the test dataset.

Precision

In terms of precision, the MLPC led with a perfect score of 1.000, followed by Random Forest at 0.714, and Logistic Regression at 0.548. These results suggest that MLPC was particularly adept at identifying true positive cases without false positives.

Recall rate

The Gaussian Naive Bayes (gnb) and AdaBoost (adab) algorithms were the top performers in the recall rate category, with scores of 0.337 and 0.198, respectively. This indicates that these algorithms were effective in identifying a significant number of true positive cases.

F1 score

The F1 score was highest for the Gaussian Naive Bayes (gnb) algorithm at 0.244, indicating a good balance between precision and recall in the test group.

AUC

Lastly, the AUC scores were highest for Logistic Regression, MLPC, and XGB, with scores of 0.669, 0.669, and 0.652, respectively. These scores reflect the algorithms' overall ability to discriminate between patients who developed POD and those who did not in the test dataset.

In terms of comprehensive performance, MLPC showed the best performance in predicting POD among these intelligent algorithms.

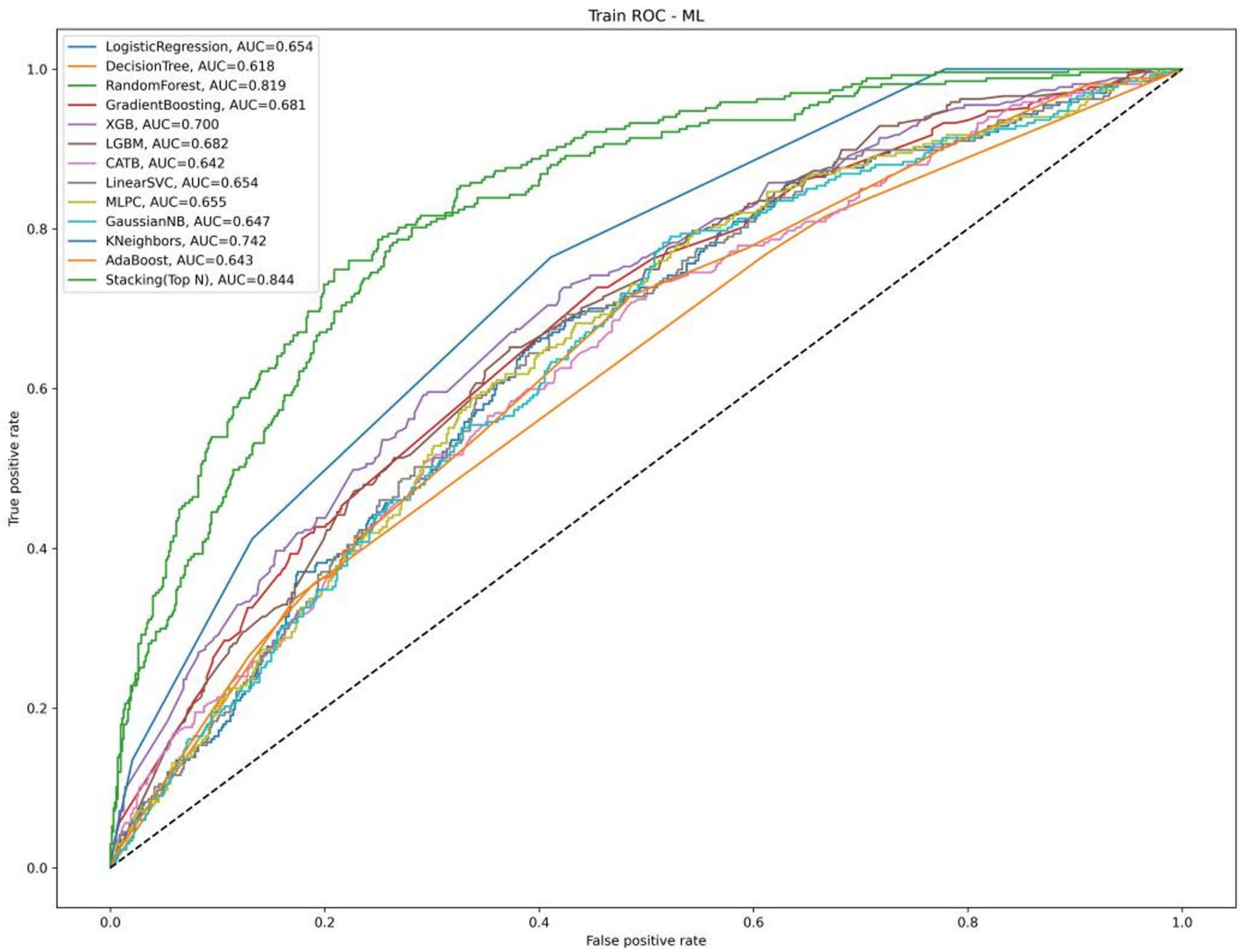

Fig. 3 Different machine learning and deep learning algorithms predict the POD in the training group Abbreviate: Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree- Gradient Boosting, light gradient boosting machine-LGBM, Extreme gradient boosting-XGB, Multilayer Perceptron Classifier-MLPC, Gaussian naive Bayes-gnb, K-nearst neighbors-knn, AdaBoost-adab

Table 3 Forecast results for testing group

Model name	AUC	Accuracy	Precision	Recall	F1
Logistic Regression	0.669	0.810	0.548	0.099	0.167
Decision Tree	0.615	0.807	0.000	0.000	0.000
Random Forest	0.591	0.810	0.714	0.029	0.056
Gradient Boosting	0.631	0.807	0.000	0.000	0.000
XGB	0.652	0.807	0.000	0.000	0.000
LGBM	0.635	0.807	0.000	0.000	0.000
MLPC	0.669	0.808	1.000	0.006	0.012
gnb	0.629	0.755	0.323	0.244	0.278
knn	0.584	0.789	0.340	0.099	0.153
adab	0.651	0.807	0.000	0.000	0.000

Abbreviate: Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree- Gradient Boosting, *LGBM* light gradient boosting machine, *XGB* Extreme gradient boosting, *MLPC* Multilayer Perceptron Classifier, *gnb* Gaussian naive Bayes, *knn* K-nearst neighbors, *adab* AdaBoost

Discussion

With the aggravation of the aging society in our country, the incidence of femoral fracture, as a common osteoporotic fracture in older patients, is gradually increasing. Older patients with intertrochanteric fracture of femur are often complicated with the medical diseases, and the risk of postoperative complications or even death is higher. Delirium is a common perioperative complication in older patients with intertrochanteric fracture of the femur, which can affect the postoperative recovery, prolong the hospitalization time, and increase the cost of treatment, and even cause the death of the patients within one year after operation [13]. Therefore, it is of great significance to evaluate delirium in older patients with femoral fracture and establish a predictive clinical decision-making system to accelerate postoperative recovery, reduce the risk of death, and improve their quality of life. The results of this study showed that the machine learning algorithm MLPC had the best

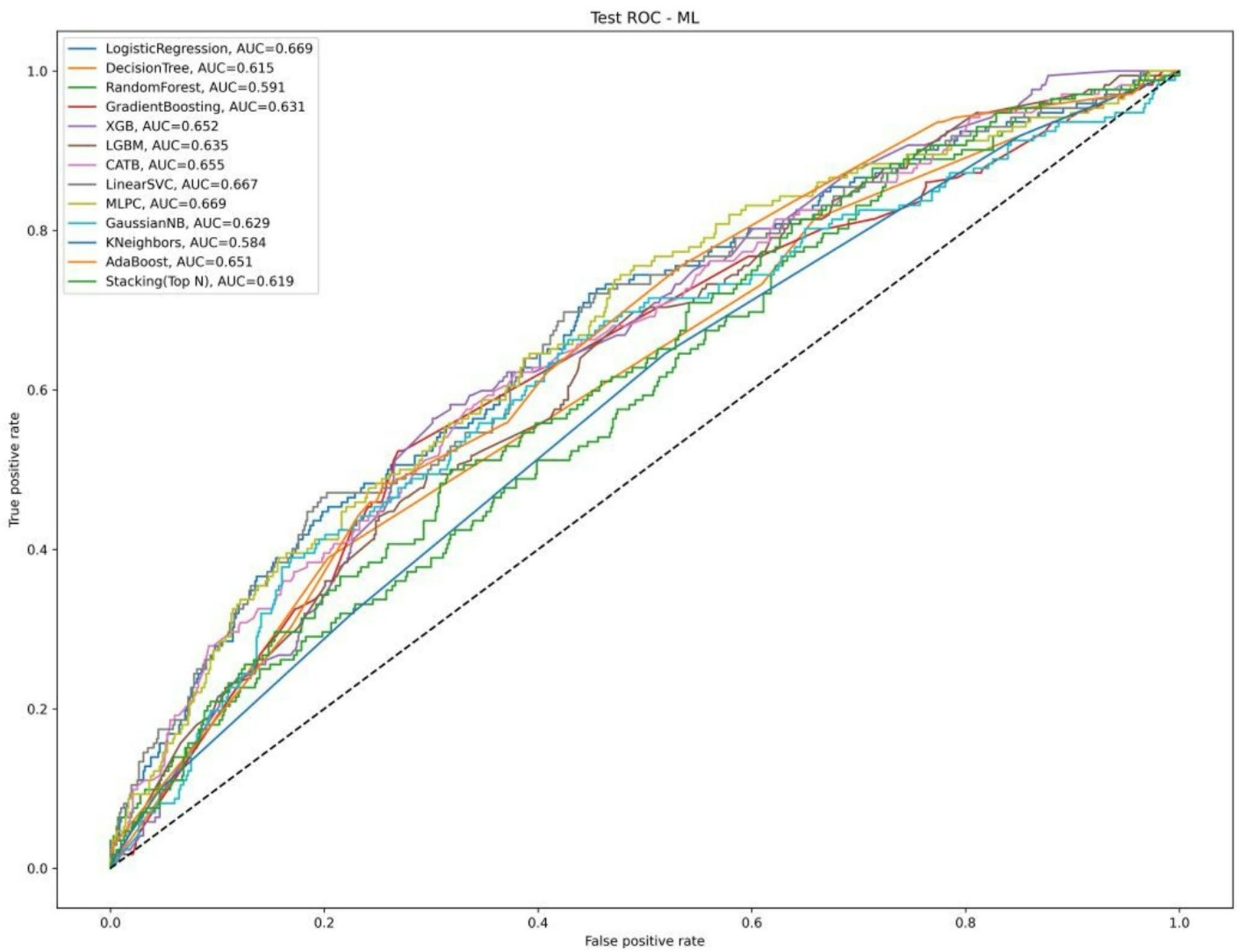

Fig. 4 Different machine learning and deep learning algorithms predict the POD in the test group. Abbreviate:Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree- Gradient Boosting, light gradient boosting machine-LGBM, Extreme gradient boosting-XGB, Multilayer Perceptron Classifier-MLPC, Gaussian naive Bayes-gnb, K-nearst neighbors-knn, AdaBoost-adab

performance in predicting POD after femoral fracture surgery in the older, with the accuracy rate of 80.8% and the precision rate of 100%. Moreover, the characteristic engineering results of GBDT algorithm revealed that age, preoperative hemoglobin, anesthesia time and intraoperative blood loss were the main factors of POD in older patients with femoral fracture.

Several studies have indicated that age, preoperative hemoglobin, anesthesia time, and intraoperative blood loss are closely related to POD. A meta-analysis including 44 studies and 104,572 subjects undergoing hip fracture surgery found that the incidence of POD was 16.93%, and 14 risk factors including age, smoking, hypertension, and chronic obstructive pulmonary disease were identified [14]. Moreover, age-adjusted Charlson comorbidity index, preoperative MMSE score, serum albumin, and postoperative pain were independently correlated with POD in older patients after thoracoabdominal surgery [15]. Multivariate logistic analysis exhibited that dementia and age ≥ 75 years old were independent risk factors for POD [16]. A meta-analysis included 15 studies discovered that nine factors were found to be related to POD, including advanced age, dementia, hypertension, use of sedative hypnotics, and low preoperative hemoglobin levels [17]. Logistic regression analysis displayed that hemoglobin, intracranial surgery, and arterial catheterization were independent risk factors for POD patients in neurosurgical intensive care unit [18]. Advanced age, dementia, diabetes, kidney disease, blood transfusion, and sedation during anesthesia recovery have been found to be associated with POD [19]. Also, it has been found that there is a close correlation between advanced age, congestive heart failure, chronic lung disease, prolonged operation time, blood transfusion and POD [20]. Moreover, central nervous system diseases, age, and blood loss are potential risk factors for POD [21]. It has been discovered that POD is closely related to smoking, prolonged operation time (more than 700 min), and blood transfusion [22]. Our research also supports this theory.

The Multilayer Perceptron Classifier (MLPC) demonstrates a certain level of practicality in the prediction of Postoperative Delirium (POD) following femoral neck surgery, particularly in dealing with high-dimensional data and capturing complex nonlinear relationships. Compared to decision tree-based algorithms such as Random Forest, MLPC may be more effective in handling nonlinear issues. In comparison to Support Vector Machines (SVM), MLPC exhibits greater flexibility in high-dimensional feature spaces. Compared to the k-Nearest Neighbors (k-NN) algorithm, MLPC is more efficient with large-scale datasets as it does not require storing training samples. Although Gradient Boosting Machines (GBM) may have higher predictive accuracy in some cases, MLPC may excel in dealing with complex interactions between features.

Our research has several potential limitations. First, this is just a model based on two data sets. Second, it does not consider the changes of related genes in the mechanism of delirium, and the next step of the study will include more older patients with femoral fracture and synthesize other feasible blood genetic indicators. Furthermore, the algorithm model included in this study is only based on the older patients with femoral fracture and cannot be extended to the prediction of delirium after orthopedic surgery in other older patients. In this study, one group was randomly selected from the 50% cross-validation, and the AUC confidence interval was not calculated by repeated sampling, so it is necessary to expand the validation times to evaluate the stability of the model. Moreover, the elderly patients included in this study have differences in basic diseases and medication, which may affect the prediction efficiency of the model for different subgroups [23]. In the future, it is necessary to analyze the subgroups of complications such as stratification and age segmentation, and make clear the applicability of the model in specific populations. For modifiable risk factors, we can verify the causal relationship between POD and target trial emulation in the future, combined with propensity score matching or instrumental variable method, so as to provide more direct evidence support for clinical intervention (such as adjusting opioid dosage) [24].

In conclusion, this study suggests that the machine learning algorithm MLPC has the best performance in predicting POD in the older patients. Age, preoperative hemoglobin, anesthesia time, and intraoperative bleeding are the most important factors for the occurrence of POD. In the future, it is of great significance to improve the prognosis and quality of life by using new intelligent method to identify, monitor and predict the risk factors of POD in older patients with femoral fracture, thus effectively preventing the perioperative delirium, lowering the risk of operation, and reducing the consumption of clinical resources.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12883-025-04389-w.

Supplementary Material 1: The study's design flowchart.

Acknowledgements
We are also grateful to the BioStudies public database for including the original data(https://www.ebi.ac.uk/biostudies/europepmc/studies/S-EPMC8766508) [25].

Authors' contributions
Y.Z., Z.N., R.L., Y.J.J. and C.M.Z. wrote the main manuscript. Z.N. and Y.Z. prepared Figs. 1, 2, 3 and 4. All authors reviewed the manuscript.

Funding
This work was supported by Guangdong Medical Research Fund (B2023054).

Data availability
We are also grateful to the BioStudies public database for including the original data(https://www.ebi.ac.uk/biostudies/europepmc/studies/S-EPMC8766508).

Declarations

Ethics approval and consent to participate
The patient information comes from the public Biostudies database. As the original data contains anonymized personal information of patients and our study is a secondary analysis of previously collected data. Using public databases for secondary analysis does not require ethical approval or informed consent, as per Article 32 of China's National Health Commission's Document No. 2023/4. The Ethics Committee of Zhanjiang Central People's Hospital also granted an exemption for the ethical review of this study, as it is a secondary analysis research involving publicly accessible databases.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Anaesthesiology and Nuring, Central People's Hospital of Zhanjiang, Zhanjiang, Guangdong, China
[2]Department of Emergency, Central People's Hospital of Zhanjiang, Guangdong, China
[3]Department of Anesthesiology, Pain and Perioperative Medicine, The first Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Received: 10 March 2024 / Accepted: 7 August 2025
Published online: 02 September 2025

References

1. Yang Y, Zhao X, Dong T, Yang Z, Zhang Q, Zhang Y. Risk factors for postoperative delirium following hip fracture repair in elderly patients: a systematic review and meta-analysis. Aging Clin Exp Res. 2017;29(2):115–26. https://doi.org/10.1007/s40520-016-0541-6.
2. Chu Z, Wu Y, Dai X, Zhang C, He Q. The risk factors of postoperative delirium in general anesthesia patients with hip fracture: attention needed. Medicine. 2021;100(22): e26156. https://doi.org/10.1097/MD.0000000000026156.
3. Jones CA, Jhangri GS, Feeny DH, Beaupre LA. Cognitive status at hospital admission: postoperative trajectory of functional recovery for hip fracture. J

Gerontol Biol Sci Med Sci. 2017;72(1):61–7. https://doi.org/10.1093/gerona/glv138.
4. Greene NH, Attix DK, Weldon BC, Smith PJ, McDonagh DL, Monk TG. Measures of executive function and depression identify patients at risk for postoperative delirium. Anesthesiology. 2009;110(4):788–95. https://doi.org/10.1097/aln.0b013e31819b5ba6.
5. Masamitsu K, Satoshi, Yurugi, Kumi M et al. Postoperative delirium in plastic or dermatologic surgery. Eur J Plast Surg. 2008;31(4):171-4. https://doi.org/10.1007/s00238-008-0245-z.
6. Siddiqi N, House AO, Holmes JD. Occurrence and outcome of delirium in medical in-patients: a systematic literature review. Age Ageing. 2006;35(4):350–64. https://doi.org/10.1093/ageing/afl005.
7. Wang Y, Lei L, Ji M, Tong J, Zhou CM, Yang JJ. Predicting postoperative delirium after microvascular decompression surgery with machine learning. J Clin Anesth. 2020;66: 109896. https://doi.org/10.1016/j.jclinane.2020.109896.
8. Zhou CM, Li H, Xue Q, Yang JJ, Zhu Y. Artificial intelligence algorithms for predicting post-operative ileus after laparoscopic surgery. Heliyon. 2024;22(5): e26580. https://doi.org/10.1016/j.heliyon.2024.e26580.
9. Xue Q, Zhu Y, Yang L, Duan W, Li Z, Ji M, et al. Predicting intraoperative bleeding in patients undergoing a hepatectomy using multiple machine learning and deep learning techniques. J Clin Anesth. 2021;74: 110444. https://doi.org/10.1016/j.jclinane.2021.110444.
10. Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). JAMA. 2001;286(21):2703–10. https://doi.org/10.1001/jama.286.21.2703.
11. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med. 1990;113(12):941–8. https://doi.org/10.7326/0003-4819-113-12-941.
12. Folstein M, Folstein S, Mchugh P, Folstein M, Folstein S, Mchugh P, et al. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189–98. https://doi.org/10.1016/0022-3956(75)90026-6.
13. Bai J, Liang Y, Zhang P, Liang X, He J, Wang J, et al. Association between postoperative delirium and mortality in elderly patients undergoing hip fractures surgery: a meta-analysis. Osteoporos Int. 2020;31(2):317–26. https://doi.org/10.1007/s00198-019-05172-7.
14. Wu J, Yin Y, Jin M, Li B. The risk factors for postoperative delirium in adult patients after hip fracture surgery: a systematic review and meta-analysis. Int J Geriatr Psychiatry. 2021;36(1):3–14. https://doi.org/10.1002/gps.5408.
15. Liu J, Li J, He J, Zhang H, Liu M, Rong J. The age-adjusted Charlson comorbidity index predicts post-operative delirium in the elderly following thoracic and abdominal surgery: a prospective observational cohort study. Front Aging Neurosci. 2022;14:979119. https://doi.org/10.3389/fnagi.2022.979119. (**Published 2022 Aug 17**).
16. Iamaroon A, Wongviriyawong T, Sura-Arunsumrit P, Wiwatnodom N, Rewuri N, Chaiwat O. Incidence of and risk factors for postoperative delirium in older adult patients undergoing noncardiac surgery: a prospective study. BMC Geriatr. 2020;20(1):40. https://doi.org/10.1186/s12877-020-1449-8. Published 2020 Feb 3.
17. Zhou Q, Zhou X, Zhang Y, Hou M, Tian X, Yang H, et al. Predictors of postoperative delirium in elderly patients following total hip and knee arthroplasty: a systematic review and meta-analysis. BMC Musculoskelet Disord. 2021;22(1): 945. https://doi.org/10.1186/s12891-021-04825-1.
18. Kose G, Şirin K, Inel MB, Mertoglu S, Aksakal R, Kurucu Ş. Prevalence and factors affecting postoperative delirium in a neurosurgical intensive care unit. J Neurosci Nurs. 2021;53(4):177–82. https://doi.org/10.1097/JNN.0000000000000595.
19. Huang J, Sprung J, Weingarten TN. Delirium following total joint replacement surgery. Bosn J Basic Med Sci. 2019;19(1):81–5. https://doi.org/10.17305/bjbms.2018.3653.
20. Romanauski TR, Martin EE, Sprung J, Martin DP, Schroeder DR, Weingarten TN. Delirium in postoperative patients admitted to the intensive care unit. Am Surg. 2018;84(6):875–80.
21. Gao H, Ma HJ, Li YJ, Yin C, Li Z. Prevalence and risk factors of postoperative delirium after spinal surgery: a meta-analysis. J Orthop Surg Res. 2020;15(1):138. https://doi.org/10.1186/s13018-020-01651-4. Published 2020 Apr 9.
22. Crawford JE, Zubair F, Baniulyte G, Wales CJ, Ansell M, Thomson E, et al. Postoperative delirium in patients with head and neck oral cancer in the West of Scotland. Br J Oral Maxillofac Surg. 2021;59(3):353–61. https://doi.org/10.1016/j.bjoms.2020.08.116.
23. Yang J, Zhang B, Chaomin H, Jiang X, Shui P, Huang J, Huang Y, Hong H, Ni Z, Zhang, et al. Identification of clinical subphenotypes of sepsis after laparoscopic surgery. Laparosc Endoscopic Robotic Surg. 2024;7:16–26. https://doi.org/10.1016/j.lers.2024.02.001.
24. Yang J, Wang L, Chen L, Zhou P, Yang S, Shen H, et al., et al. A comprehensive step-by-step approach for the implementation of target trial emulation: evaluating fluid resuscitation strategies in post-laparoscopic septic shock as an example. Laparosc Endoscopic Robotic Surg. 2025;8(1):28–44.
25. Liu YM, Huang H, Gao J, Zhou J, Chu HC. Hemoglobin concentration and post-operative delirium in elderly patients undergoing femoral neck fracture surgery. Front Med. 2022;8: 780196. https://doi.org/10.3389/fmed.2021.780196.

Publisher's Note