Clinical Interventions in Aging

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Machine Learning-Based Prediction of Postoperative Pneumonia Among Super-Aged Patients With Hip Fracture

Miaotian Tang[1], Meng Zhang[1], Yu Dang[1], Mingxing Lei[2], Dianying Zhang[1,3–5]

[1]Department of Trauma Orthopaedics, Peking University People's Hospital, Beijing, 100044, People's Republic of China; [2]Department of Orthopaedics, Hainan Hospital of Chinese PLA General Hospital, Sanya, 572013, People's Republic of China; [3]National Trauma Medical Center, Beijing, 100044, People's Republic of China; [4]Key Laboratory of Trauma Treatment and Neural Regeneration, Ministry of Education, Beijing, 100044, People's Republic of China; [5]Department of Orthopedics, Peking University Binhai Hospital, Tianjin, 300450, People's Republic of China

Correspondence: Dianying Zhang, Department of Trauma Orthopaedics, Peking University People's Hospital, Beijing, 100044, People's Republic of China, Tel +86 010-88326550, Email zdy8016@163.com; Mingxing Lei, Department of Orthopaedics, Hainan Hospital of Chinese PLA General Hospital, Sanya, 572013, People's Republic of China, Tel +8618811772189, Email leimingxing2@sina.com

Background: Hip fractures have become a significant health concern, particularly among super-aged patients, who were at a high risk of postoperative pneumonia due to their frailty and the presence of multiple comorbidities. This study aims to establish and validate a model to predict postoperative pneumonia among super-aged patients with hip fracture.

Methods: Data were derived from the Chinese PLA General Hospital (PLAGH) Hip Fracture Cohort Study, and we included 555 super-aged patients (≧80 years old) with hip fracture treated with surgery. Patient's demographics, comorbidities, laboratory tests, and surgery types were collected for analysis. All patients were randomly splitting into a training group and a validation group according to the ratio of 7:3. The majority of patients were used to train models, which was tuned using a series of algorithms, including decision tree (DT), random forest (RF), extreme gradient boosting machine (eXGBM), support vector machine (SVM), neural network (NN), and logistic regression (LR).

Results: The incidence of postoperative pneumonia was 7.2% (40/555). Among the six developed models, the eXGBM model demonstrated the optimal model, with the area under the curve (AUC) value of 0.929 (95% CI: 0.900–0.959), followed by the RF model (AUC: 0.916, 95% CI: 0.885–0.948). The LR model had an AUC value of 0.720 (95% CI: 0.662–0.778). In addition, the eXGBM model demonstrated the optimal prediction performance in terms of accuracy (0.858), precision (0.870), F1 score (0.855), Brier score (0.104), and log loss (0.349). It also showed favorable calibration ability and favorable clinical net benefits across various threshold risk.

Conclusion: This study develops and validates a reliable machine learning-based model to predict pneumonia specifically among super-aged patients with hip fracture following surgery. This model can serve as a useful tool to identify postoperative pneumonia and guide clinical strategies for super-aged patients with hip fracture.

Keywords: machine learning, postoperative pneumonia, hip fracture, super-aged patients, geriatric patients

Introduction

With the global aging population on the rise, hip fractures have become a significant health concern,[1] particularly among the older adults. Super-aged patients, defined as those aged 80 years and older,[2,3] were at a heightened risk of complications following surgery due to their frailty and the presence of multiple comorbidities. One of the most severe and common complications in this demographic is postoperative pneumonia,[4] which can lead to increased morbidity, extended hospital stays, and higher mortality rates. Given the aging population and the increasing prevalence of hip fractures, addressing postoperative pneumonia was a critical component of improving surgical outcomes and overall patient care among super-aged patients with hip fracture.

Received: 18 November 2024
Accepted: 19 February 2025
Published: 27 February 2025

The proposal of predictive models relied on risk factors, and numerous studies have already reported risk factors associated with postoperative pneumonia following hip fractures.[5–8] The main factors included age,[5–8] ASA score,[5,7] smoking,[5] diabetes,[5] coronary heart disease,[5,6,8] chronic obstructive pulmonary disease,[5] dementia,[5,6] hemoglobin levels,[7] albumin levels,[7] and blood urea nitrogen.[7] These factors provided significant guidance for clinical practice in preventing postoperative lung infections among patients with hip fracture. Notably, using well-established risk factors to build a model could increase its robustness and enhance its predictive performance.

Thus, several prediction models have been specifically developed for patients with hip fractures to assess post-operative pneumonia,[9,10] acute respiratory failure,[11] and early infections.[12] These models can be beneficial for clinical guidance in preventing postoperative pneumonia in patients with hip fractures. However, the AUC values of these models were mostly below 0.85, indicating that their accuracy needs further improvement. Traditional clinical assessment tools might not be sufficiently predictive for this high-risk group, thus necessitating the development of more advanced and precise predictive models. Recent advancements in machine learning offer a promising avenue for enhancing predictive accuracy in medical outcomes.[13,14] Machine learning algorithms could analyze vast amounts of data and identify complex patterns that might be overlooked by conventional methods.[15,16] By incorporating various patient-specific factors, these algorithms could provide personalized risk assessments, thereby aiding clinicians in identifying patients at high risk for postoperative pneumonia and tailoring preventive strategies accordingly. For instance, there were two prediction models for postoperative pneumonia in hip fracture patients,[17,18] and the two studies employed different statistical and machine learning approaches. Nonetheless, previous studies primarily included populations aged 60 or 65 years and above, rather than super-aged patients.

Therefore, the objective of this study was to establish and validate a robust machine learning-based model to predict the occurrence of postoperative pneumonia in super-aged patients undergoing surgery for hip fracture. By accurately identifying those at higher risk, healthcare providers could implement targeted interventions and preventive measures, ultimately reducing the incidence of this serious complication, enhancing patient recovery, and optimizing the use of medical resources.

Patients and Methods

Patients and Study Design

The study reanalyzed data of 555 patients from the Chinese PLA General Hospital (PLAGH) Hip Fracture Cohort Study. The study focused on hip fracture patients diagnosed with femoral neck fractures or femoral intertrochanteric fractures. All patients meeting the following criteria were included: (1) diagnosed with a femoral neck or intertrochanteric fracture, and (2) received surgical treatment. Exclusion criteria were: (1) patients under 80 years old, (2) radiograph-confirmed pneumonia within 48 hours of admission, (3) patients receiving conservative treatment, and (4) history of pneumonia within 3 months prior to hospitalization. Figure 1 illustrates the patient enrollment flowchart. The Medical Research Ethics Board of the Chinese PLA General Hospital approved the study protocol and waived the need for patient consent for the review of medical records and images due to the retrospective nature of the review. In addition, the data was anonymized or maintained with confidentiality. The study was conducted in accordance with the Declaration of Helsinki. In addition, the study adhered to the STROCSS criteria[19] and the TRIPOD Checklist.[20]

Collection of Variables

This study gathered a comprehensive range of variables, encompassing baseline characteristics like age, gender, body mass index (BMI), and smoking status. It also included data on comorbidities such as chronic obstructive pulmonary disease (COPD), diabetes, hypertension, coronary heart disease, previous stroke, dementia, digestive system disorders, renal disease, and the total number of comorbidities. Additionally, information on fractures, including fracture type, and surgical details such as the American Society of Anesthesiologists (ASA) classification, type of surgery, and use of mechanical ventilation, were collected. Laboratory test results, including hemoglobin (HB), albumin levels, blood creatinine, and red cell distribution width (RDW), were also documented. All these variables were recorded post-hospitalization but prior to surgery.

 https://doi.org/10.2147/CIA.S507138

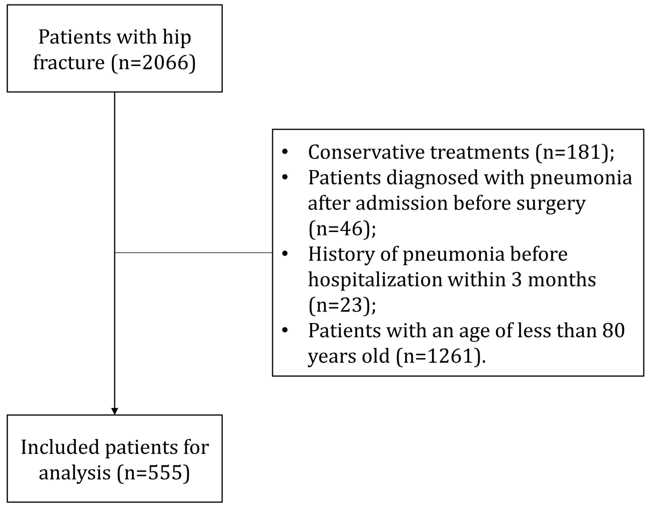

Figure 1 Patient's flowchart.

Identification of Model Features

Feature importance analysis was utilized to assess the relative significance of each feature in influencing the outcome. By pinpointing the most impactful features, we gained insights into the key factors driving the results, thereby enhancing the interpretability and validation of the model's predictions. In this investigation, Shapley Additive Explanation (SHAP) values were employed to evaluate the importance of each input parameter.[21,22] SHAP values specifically measure the contribution of each feature to the model's output. For this study, variables with SHAP values equal to or greater than 0.25 were deemed significant model features.

Definition of Outcome

Postoperative pneumonia was diagnosed based on the presence of new infiltrates on chest X-ray, along with one or more of the following symptoms observed during hospitalization after surgery: (1) new or worsening respiratory symptoms such as coughing and expectoration; (2) fever or hypothermia; (3) physical examination revealing signs of lung consolidation or moist rales; (4) a white blood cell count exceeding 10×10^9/L or below 4×10^9/L; (5) isolation of pathogens from blood culture or sputum.

Modeling

This study employed six algorithms to train and optimize models: Decision Tree (DT), eXtreme Gradient Boosting Machine (eXGBM), Support Vector Machine (SVM), Random Forest (RF), Neural Network (NN), and Logistic Regression (LR). Each model was provided with the same input features to maintain consistency. Prior to modeling, the SMOTETomek resampling strategy[23–25] was applied to address data imbalance and enhance model robustness. A comprehensive data preprocessing pipeline ensured consistent and reproducible data transformations. Additionally, a stratified sampling approach was used to preserve the proportion of outcome classes within sub-datasets. During modeling, both grid and random hyperparameter searches were conducted alongside cross-validation to determine the optimal hyperparameters for each model. The Area Under the Curve (AUC) was used as the primary optimization metric. To accommodate variability in model performance, a wide range of hyperparameter values was explored, resulting in a combination of underfitted and overfitted models.

Model Validation

The models were validated using several evaluation metrics, such as AUC, accuracy, precision, recall, F1 score, Matthews's correlation coefficient (MCC), Brier score, and log loss. To evaluate the models' discriminative ability, calibration, and clinical net benefits, probability density curves, calibration curves, and decision curves were employed, respectively. Additionally, a confusion matrix was utilized to assess accuracy, precision, and recall metrics.

https://doi.org/10.2147/CIA.S507138

The Brier Score is a metric used to evaluate the accuracy of probabilistic predictions. It quantifies the mean squared difference between the predicted probabilities and the actual binary outcomes, providing a clear measure of how closely the predictions align with the observed results. It is defined mathematically as:

$$Brier\ Score = \frac{1}{N}\sum_{i=1}^{n}(p_i - o_i)^2$$

Where, N represents the total sample, p_i is the predicted risk of postoperative in-hospital mortality, and o_i is the actual probability of postoperative in-hospital mortality.

Log loss, commonly referred to as logistic loss or cross-entropy loss, is a widely used performance metric in classification tasks, especially in binary classification scenarios. This metric evaluates the accuracy of a classifier by comparing the predicted probabilities with the actual class labels, thereby providing insight into the model's performance. The log loss formula is:

$$log\ Loss = \frac{-1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\, log\left(p_{ij}\right)$$

Where, N is the number of samples, M is the number of classes, y_{ij} represents the true label of sample i for class j (0 or 1), and p_{ij} is the predicted probability of sample i belonging to class j.

Statistical Analysis

Categorical variables were represented as percentages, while continuous data were summarized using medians and interquartile ranges. Machine learning algorithms and hyperparameter tuning were conducted using Python (version 3.9.7) with the scikit-learn library (version 1.2.2). All other statistical analyses were performed using R (version 4.1.2). A p-value of less than 0.05 was deemed statistically significant.

Results

Baseline Characteristics

The baseline clinical characteristics of 555 patients who underwent surgery for hip fractures were summarized in Table 1. The median age was 84 years (IQR 82–88), with 32.1% male and 67.9% female patients. The median BMI was 22.49 kg/m² (IQR 19.89–23.41). The fracture types included 37.3% intertrochanteric and 62.7% femoral neck fractures. The

Table 1 Baseline Clinical Characteristics Among Patients With Hip Fracture Treated With Surgery

Characteristics	Overall
n	555
Age (years, median [IQR])	84.00 [82.00, 88.00]
Gender (male/female, %)	178/377 (32.1/67.9)
BMI (kg/m^2, median [IQR])	22.49 [19.89, 23.41]
Fracture type (intertrochanteric/femoral neck, %)	207/348 (37.3/62.7)
Smoking (no/yes, %)	536/19 (96.6/3.4)
COPD (no/yes, %)	524/31 (94.4/5.6)
Diabetes (no/yes, %)	425/130 (76.6/23.4)
Hypertension (no/yes, %)	233/322 (42.0/58.0)

(*Continued*)

Table 1 (Continued).

Characteristics	Overall
Coronary heart disease (no/yes, %)	412/143 (74.2/25.8)
Previous stroke (no/yes, %)	464/91 (83.6/16.4)
Dementia (no/yes, %)	523/32 (94.2/5.8)
Digestive system disorders (no/yes, %)	531/24 (95.7/4.3)
Renal disease (no/yes, %)	537/18 (96.8/3.2)
ASA (%)	
1	11 (2.0)
2	227 (40.9)
3	285 (51.4)
4	32 (5.8)
Surgery type (%)	
Intramedullary fixation	231 (41.6)
Hip replacement	279 (50.3)
Other procedures	45 (8.1)
Mechanical ventilation (no/yes, %)	383/172 (69.0/31.0)
HB (g/L, median [IQR])	118.00 [103.00, 130.00]
Albumin (g/L, median [IQR])	36.30 [33.00, 39.35]
Blood creatinine (μmol/L, median [IQR])	65.40 [60.05, 76.55]
RDW (%, median [IQR])	13.30 [12.70, 14.00]
Number of comorbidities (%)	
0	148 (26.7)
1	157 (28.3)
2	115 (20.7)
≧3	135 (24.3)
Postoperative pneumonia (no/yes, %)	515/40 (92.8/7.2)

Abbreviations: IQR, Interquartile range; BMI, Body mass index; COPD, Chronic obstructive pulmonary disease; ASA, American society of anesthesiologists; HB, Hemoglobin; RDW, Red cell volume distribution width.

prevalence of smoking was low (3.4%), while other comorbid conditions such as COPD (5.6%), diabetes (23.4%), hypertension (58.0%), coronary heart disease (25.8%), and previous stroke (16.4%) were noted. Dementia was present in 5.8% of patients, and 4.3% had digestive system disorders. Renal disease was observed in 3.2% of patients. According to the ASA classification, 40.9% were ASA 2, 51.4% were ASA 3, and 5.8% were ASA 4. Regarding the type of surgery, 41.6% underwent intramedullary fixation, 50.3% had hip replacements, and 8.1% had other procedures. Mechanical ventilation was required in 31.0% of cases. Median hemoglobin level was 118 g/L (IQR 103–130), median albumin was 36.3 g/L (IQR 33.0–39.35), and median blood creatinine was 65.4 μmol/L (IQR 60.05–76.55). The median RDW was

13.3% (IQR 12.7–14.0). Comorbidities were present in varying numbers: 26.7% had no comorbidities, 28.3% had one, 20.7% had two, and 24.3% had three or more. Postoperative pneumonia occurred in 7.2% of patients.

Identification of Model Features

Based on SHAP, the variable importance ranking was as follows: RDW, blood creatinine, surgery type, HB, albumin, age, diabetes, BMI, the number of comorbidities, coronary heart disease, ASA, gender, dementia, previous stroke, fracture type, mechanical ventilation, hypertension, renal disease, COPD, digestive system disorders, and smoking (Figure 2A). There were 10 variables with SHAP values greater than or equal to 0.25: RDW, blood creatinine, surgery type, HB, albumin, age, diabetes, BMI, the number of comorbidities, and coronary heart disease (Figure 2B). The top three variables in terms of importance were RDW, blood creatinine, and surgery type, with SHAP values of 0.69, 0.42, and 0.41, respectively. Hence, these 10 variables were used as model features to train models.

Model's Prediction Performance

Among the six models, the eXGBM model demonstrated the optimal model, with the area under the curve (AUC) value of 0.929 (95% CI: 0.900–0.959), followed by the RF model (AUC: 0.916, 95% CI: 0.885–0.948). The LR model only had an AUC value of 0.720 (95% CI: 0.662–0.778) (Figure 3). In addition, the eXGBM model demonstrated the best overall prediction performance across multiple metrics. It achieved the highest values for accuracy (0.858), precision (0.870), and F1 score (0.855) (Table 2). Additionally, the eXGBM model had the best Brier score (0.104) and the lowest log loss (0.349), indicating superior calibration and overall model reliability. Regarding MCC, the eXGBM model also had the highest MCC value of 0.717, suggesting it performs best among the models evaluated. The RF model followed closely with an MCC of 0.695, while the NN model and the SVM model showed moderate performance with scores of 0.683 and 0.459, respectively. The DT model and the LR model had lower MCC values of 0.439 and 0.356, indicating relatively poorer performance compared to the other models.

The probability density curve analysis of the DT model showed that patients with postoperative pneumonia could be distinguished from those without it, but there was considerable overlap between the two groups (Figure 4A). In contrast, the eXGBM model (Figure 4B), SVM model (Figure 4C), RF model (Figure 4D), and NN model (Figure 4E) demonstrated more favorable predictive performance, as the peak risk probability for the postoperative pneumonia group shifted to the right, while that for the non-postoperative pneumonia group shifted to the left. However, the LR model exhibited the poorest

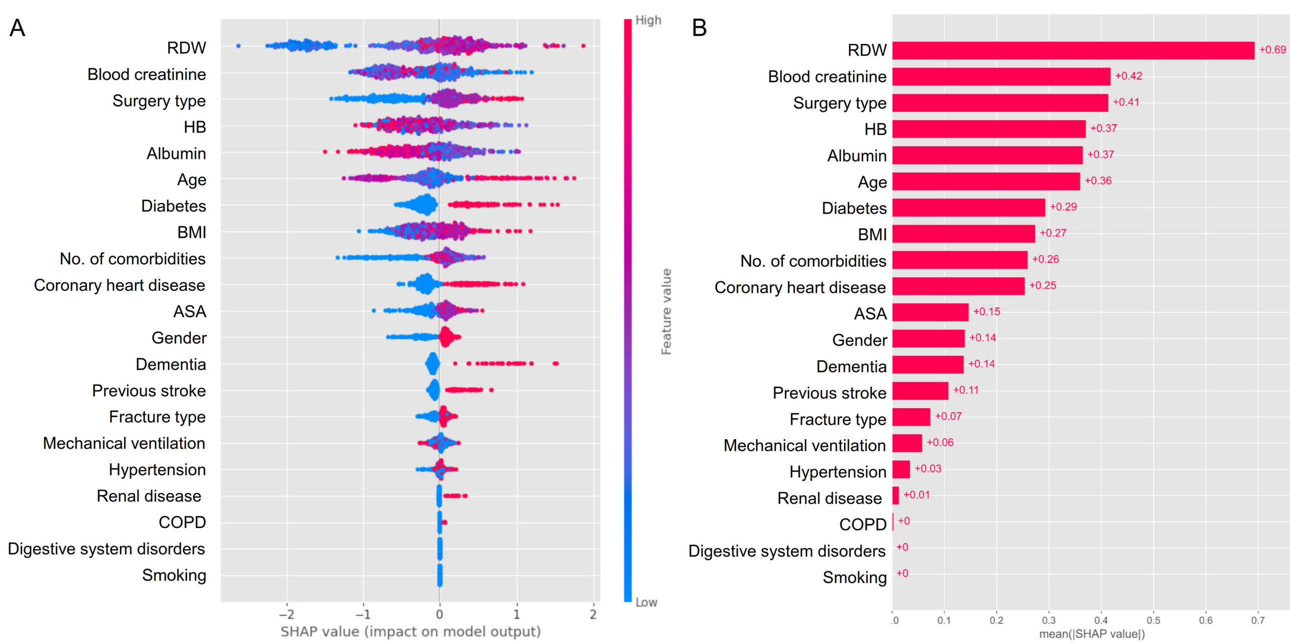

Figure 2 Feature importance analysis. (**A**) Feature importance ranking; (**B**) SHAP value of all features.

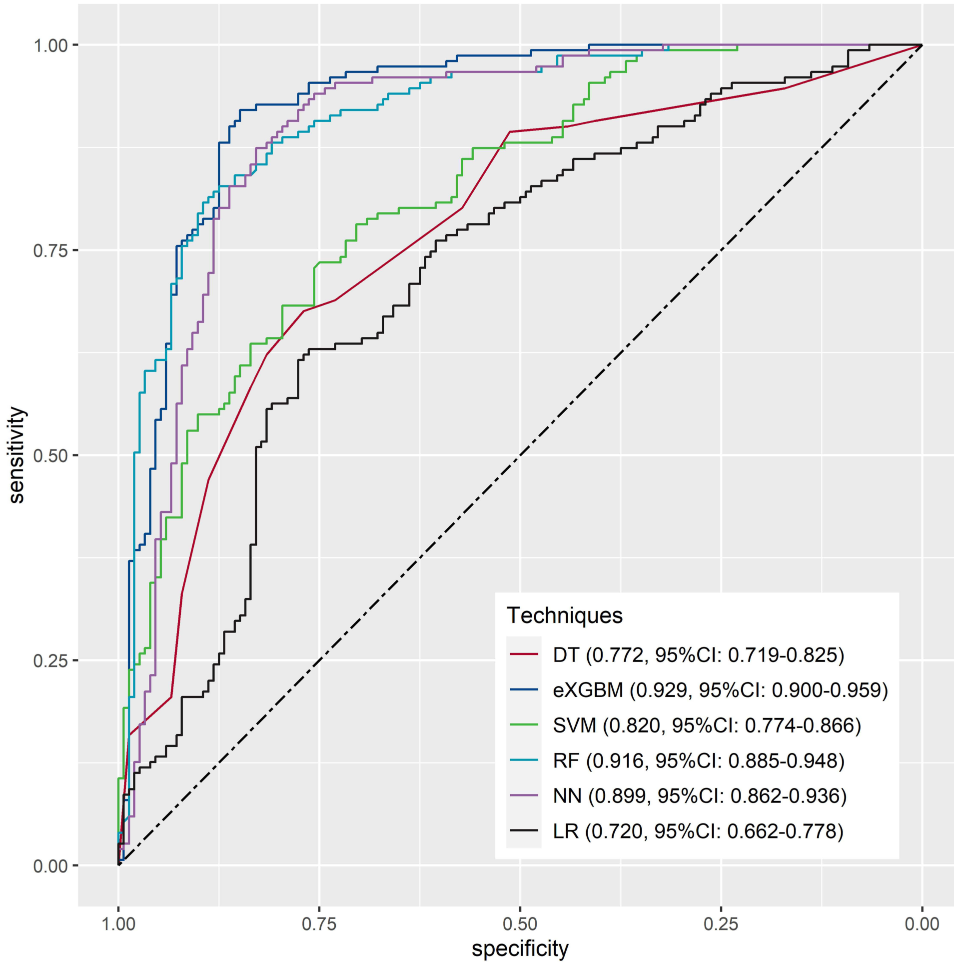

Figure 3 Area under the curve analysis for all machine learning models.
Abbreviations: DT, decision tree; eXGBM, extreme gradient boosting machine; SVM, support vector machine; RF, random forest; NN, neural network; LR, logistic regression.

discrimination (Figure 4F). Calibration curve analysis demonstrated that most models had favorable calibration ability, as their curves closely aligned with the ideal calibration curve (Figure 5). Additionally, decision curve analysis indicated that the eXGBM model, the RF model, and the NN model provided favorable clinical net benefits across various threshold risks (Figure 6).

Individualized Prediction

Individualized predictions were achieved in this study using the optimal model, the eXGBM model. For instance, in a true positive case (Figure 7), the patient was 89 years old and underwent hip replacement surgery. The patient's hemoglobin (HB) level was 97 g/L, albumin was 26.8 g/L, blood creatinine was 145 μmol/L, red cell distribution width (RDW) was 17%, and body mass index (BMI) was 22.2 kg/m². The patient had three comorbidities but did not have coronary heart disease or diabetes. In this scenario, HB, albumin, and surgery type were the most significant features contributing to the risk of postoperative pneumonia.

In another case, which was a true negative (Figure 8), the patient was 84 years old with an RDW of 12.8%, albumin of 40.9 g/L, HB of 124 g/L, blood creatinine of 78.9 μmol/L, and a BMI of 24.0 kg/m². This patient had one comorbidity, was treated with hip replacement, and did not have coronary heart disease or diabetes. The total score for this case was −6.07, which was lower than the base score of −0.02, indicating a low risk of postoperative pneumonia.

 https://doi.org/10.2147/CIA.S507138

Table 2 Prediction Performance of All Models in the Study

Metrics	Models					
	DT	eXGBM	SVM	RF	NN	LR
Accuracy	0.723	0.858	0.729	0.848	0.842	0.677
Precision	0.745	0.870	0.741	0.878	0.837	0.657
Recall	0.675	0.841	0.702	0.808	0.848	0.735
F1 score	0.708	0.855	0.721	0.841	0.842	0.694
AUC	0.772	0.929	0.820	0.916	0.899	0.720
AUC (95% CI)	0.719–0.825	0.900–0.959	0.774–0.866	0.885–0.948	0.862–0.936	0.662–0.778
MCC	0.439	0.717	0.459	0.695	0.683	0.356
Brier score	0.194	0.104	0.175	0.124	0.123	0.213
Log loss	1.434	0.349	0.522	0.400	0.406	0.614

Abbreviations: DT, decision tree; eXGBM, extreme gradient boosting machine; SVM, support vector machine; RF, random forest; NN, neural network; LR, logistic regression; AUC, area under the curve; CI, confident interval; MCC, Matthews's correlation coefficient.

Discussion

Main Findings

The main finding of this study is the development and validation of a machine learning-based model to predict postoperative pneumonia in super-aged patients with hip fractures. The eXGBM model, in particular, demonstrates superior predictive performance and calibration compared to other algorithms tested. This model identifies key variables such as RDW, blood creatinine, surgery type, and hemoglobin, providing a robust basis for accurate risk stratification.

The findings from our study indicate a promising application of machine learning techniques in the clinical setting, particularly for predicting postoperative pneumonia in super-aged patients with hip fractures. By identifying patients at elevated risk for postoperative pneumonia, healthcare providers can implement targeted preventive measures and tailored postoperative care protocols. Furthermore, the integration of this predictive model into routine clinical workflows can facilitate early interventions that may reduce the incidence of postoperative complications, ultimately improving patient outcomes and hospital resource utilization.

Incidence of Postoperative Pneumonia in Hip Fracture Patients

The incidence of postoperative pneumonia in hip fracture patients varied across different studies. Previous research has reported rates ranging from 1.8% to 9.26%,[5–7,26] depending on the population studied and the definitions used for pneumonia. A meta-analysis included 12,084 patients for analysis, and found 809 patients had postoperative pneumonia, and thus the incidence of postoperative pneumonia was found to be 6.69%.[5] Another meta-analysis revealed that the incidence of postoperative pneumonia was 7.8% after analyzing ten studies consisting of 37,130 patients.[7] In our study, the incidence of postoperative pneumonia was found to be 7.2%, which falls within the previously reported range. This comparison highlights the consistency of our findings with existing literature and underscores the clinical relevance of our predictive model.

Risk Factors for Postoperative Pneumonia in Hip Fracture Patients

Several risk factors for postoperative pneumonia in hip fracture patients have been identified in the previous literature. For example, in a meta-analysis of ten studies, Gao et al[5] demonstrated that older age, male gender, ASA score of three or above, dependent functional status, smoking, chronic obstructive pulmonary disease, diabetes mellitus, coronary heart disease, arrhythmia, cerebrovascular disease, dementia, chronic renal failure, hip arthroplasty, delayed surgery,

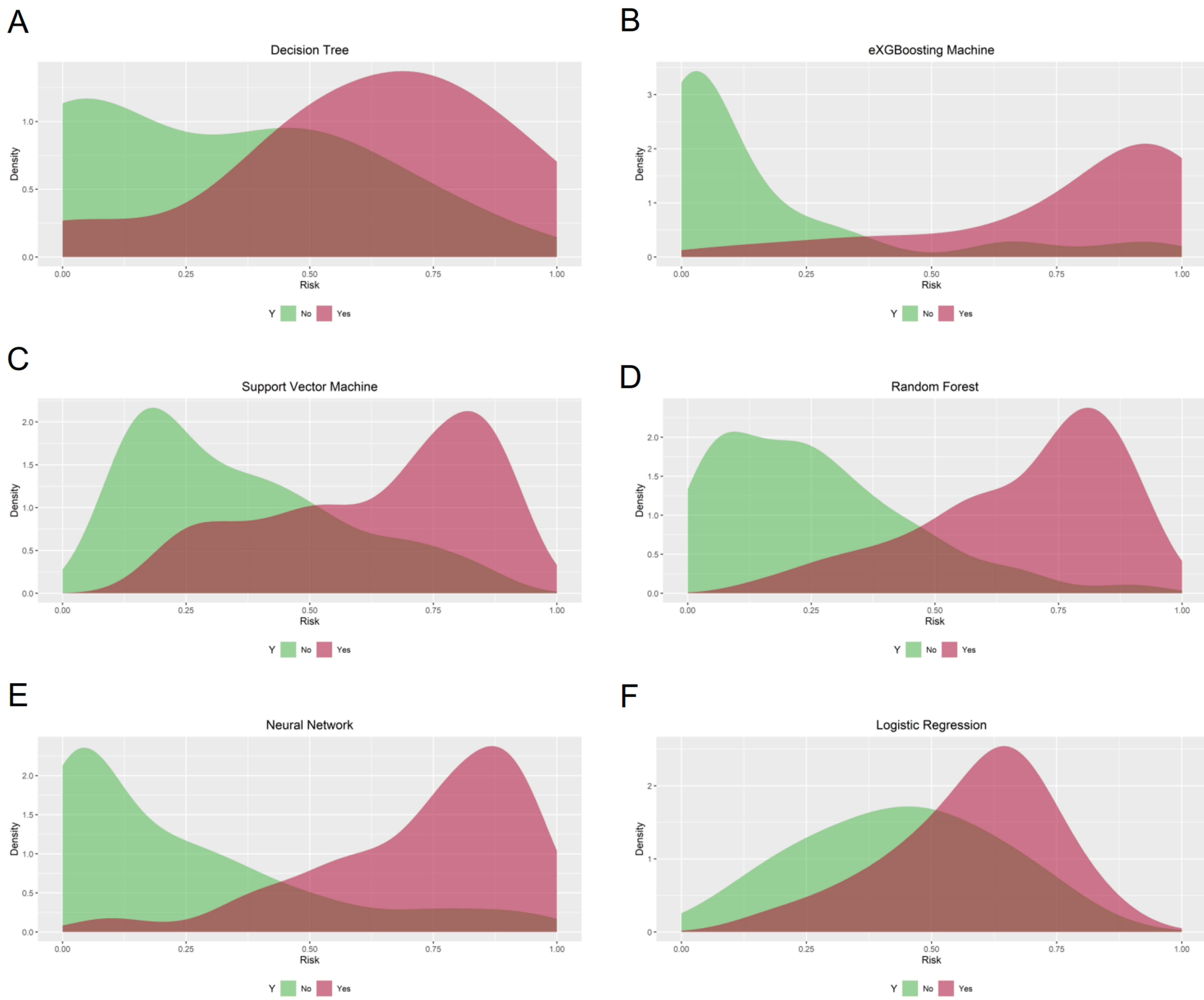

Figure 4 Probability density curve analysis for all machine learning models. (**A**) Decision tree; (**B**) Extreme gradient boosting machine; (**C**) Support vector machine; (**D**) Random forest; (**E**) Neural network; (**F**) Logistic regression.

preoperative creatinine, and preoperative serum albumin were significantly associated with postoperative pneumonia. A large retrospective cohort study found that the predictive factors of pneumonia among older patients treated with hip fracture surgery included advanced age, male sex, lean body, cerebrovascular disease, dementia, and dependency for activities of daily living.[6] A study of collecting 1208 patients aged more than 65 years and treated with hip fracture surgery also revealed that an ASA score of 3 or above, higher Charlson Comorbidity Index, and postoperative delirium were significantly risk factors for postoperative pneumonia.[26] More recently, Yao et al[27] neutrophil-to-lymphocyte ratio was an important predictor for postoperative pneumonia among older adults with hip fracture. Han et al[7] also conducted a meta-analysis of ten studies and found that advanced age, male sex, ASA score of ≧3, chronic obstructive pulmonary disease, coronary heart disease, arrhythmia, congestive heart failure, chronic kidney disease, and cerebrovascular accident. Additionally, the following factors are associated with postoperative pneumonia: Hemoglobin levels, albumin levels, blood urea nitrogen, alanine aminotransferase, arterial oxygen pressure, time from injury to surgery, and surgery within 48 hours. A retrospective analysis also revealed that age, sex, respiratory disease, heart disease, cerebrovascular disease, liver disease, preoperative stay, general anesthesia was found to be risk factors for postoperative pneumonia.[8] Our findings were consistent with previous studies.[28,29] In our studies, we also found that surgery type, albumin, age, diabetes, BMI, the number of comorbidities, and coronary heart disease were the most important ten contributors to

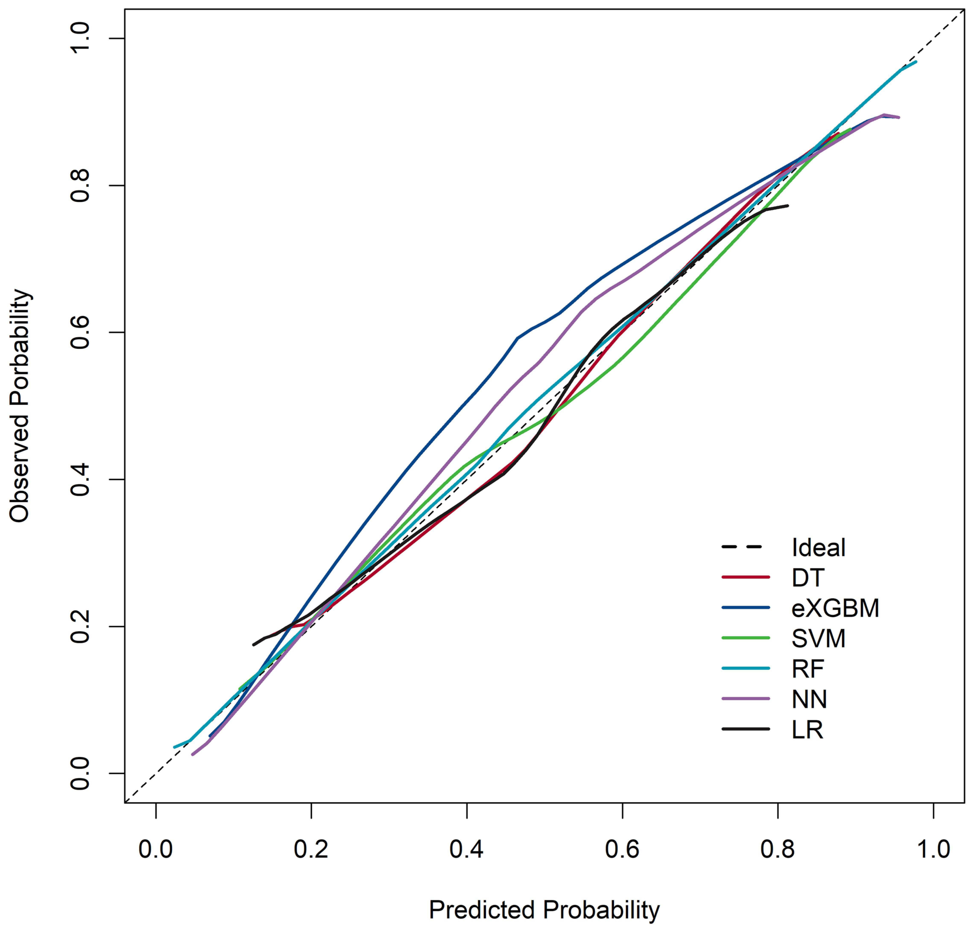

Figure 5 Calibration curve analysis for all machine learning models.

postoperative pneumonia. Additionally, our model highlights the importance of specific laboratory values like RDW and blood creatinine, which are less frequently reported in previous studies. Using well-established risk factors to build a model can increase its robustness and enhance its predictive performance. In addition, this broader inclusion of variables may enhance the predictive power of our model and provide a more comprehensive understanding of the risk factors involved.

Comparison With Previous Predictive Models

After comprehensively reviewing literature, we found that there were two prediction models for postoperative pneumonia in hip fracture patients, and the two studies employed different statistical and machine learning approaches. For example, Guo et al[17] developed machine learning models to identify older adults at high risk of postoperative pneumonia after hip fracture surgery. Conducted at a central hospital, the study included 805 patients (≧60 years), with 75 (9.3%) developing postoperative pneumonia. Data collected within 24 hours of admission were used to build models with seven machine learning algorithms: CART, GBM, KNN, LR, NNet, RF, and XGBoost. Key predictive features included age, cerebral infarction, COPD, WBC, HB, GLU, STB, GLOB, and Ka^{+}. Among these, the XGBoost model demonstrated the best performance, making it highly effective for early detection of the risk of postoperative pneumonia among older adults with fracture. Another study conducted by Dai et al[18] also developed and validated a machine learning model, specifically the Catboost model, to predict postoperative pneumonia in older adults with hip fractures. Using clinical data from two hospitals, the model demonstrated high predictive accuracy, with AUC values of 0.895 and 0.835 in training and testing sets, respectively, and 0.894 during external validation. Key predictors identified included CRP levels, the modified five-item frailty index, and ASA body status. Notably, in comparison to these models, our eXGBM-

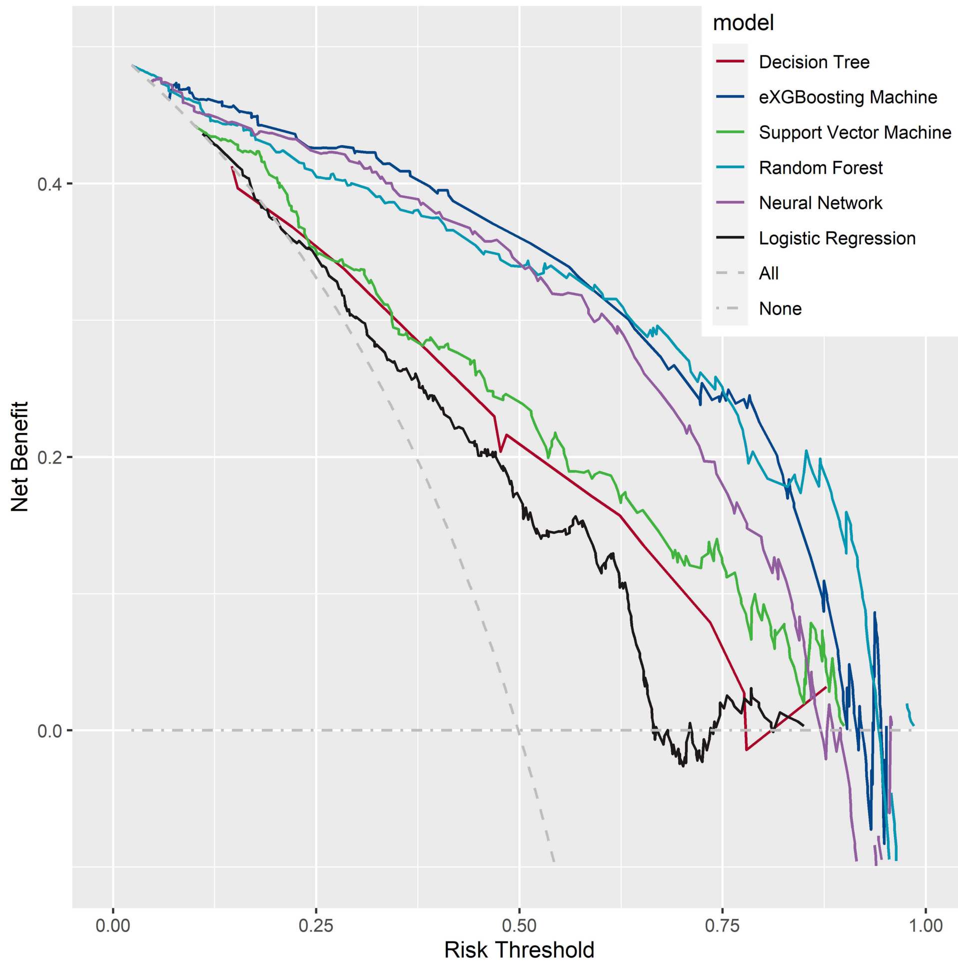

Figure 6 Decision curve analysis for all machine learning models.

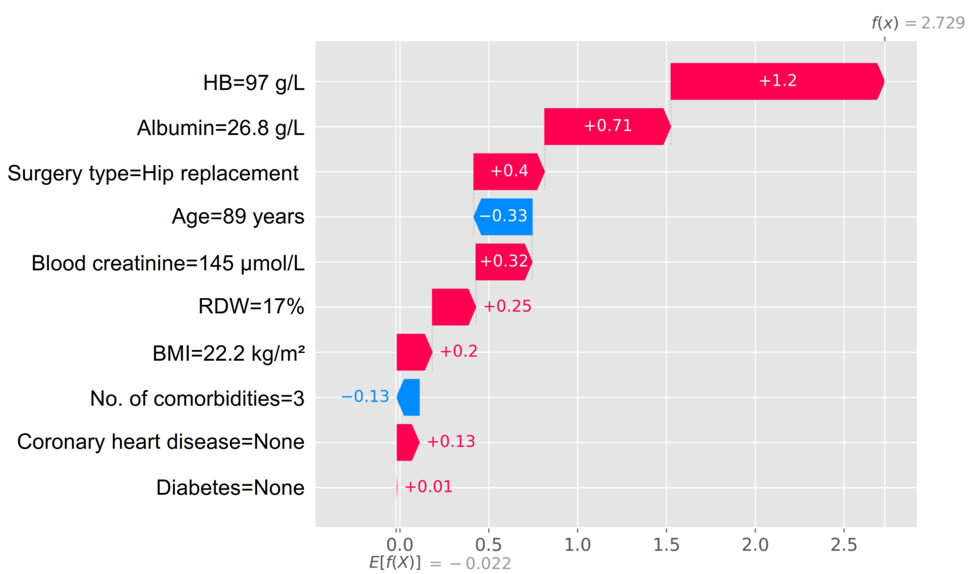

Figure 7 Individualized prediction: A true positive case.

https://doi.org/10.2147/CIA.S507138

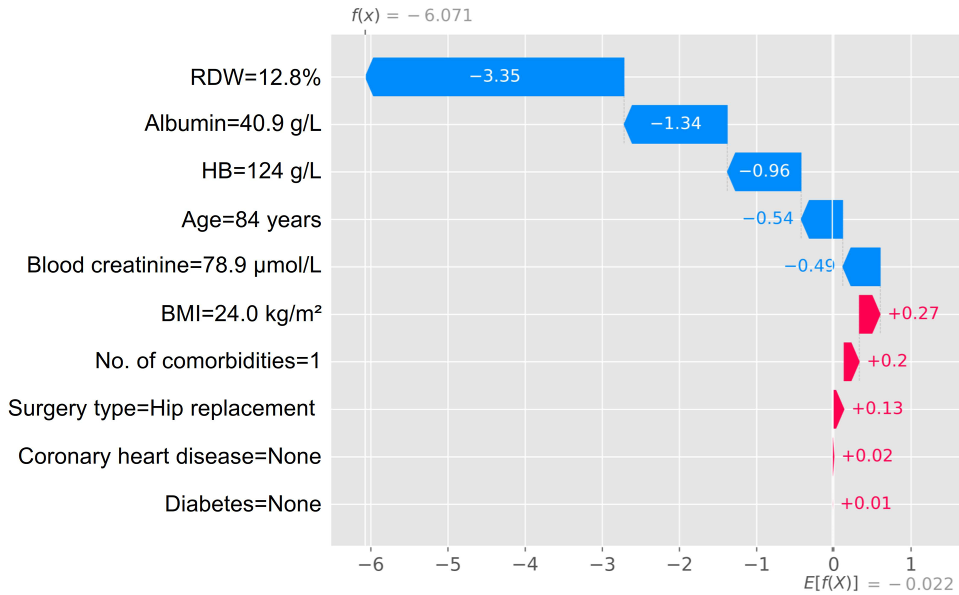

Figure 8 Individualized prediction: A true negative case.

based approach exhibited superior predictive performance, achieving an AUC value of 0.929. Additionally, the MCC for the eXGBM model was the highest among all, with a value of 0.717. Our model was developed based on a cohort of super-aged patients, specifically including those aged 80 years and older. In contrast, previous studies primarily included populations aged 60 or 65 years and above. Incorporating patients aged 80 years and older into our study allowed for a more targeted and relevant analysis for this unique demographic, who often presented different clinical characteristics and healthcare needs compared to relatively younger populations. Furthermore, it enhanced the generalizability and applicability of our model to the oldest-old population, who were typically underrepresented in clinical research. In addition, the inclusion of a wide range of variables and the application of a robust machine learning framework distinguished our model from previous efforts. However, like earlier models, our study underscored the need for further external validation to confirm its applicability across diverse clinical settings. Additionally, integrating machine learning-based models with clinical practice is the future trend, although there are still numerous challenges in the eHealth field that need to be addressed to improve patient outcomes.[30]

Limitations

Despite the promising results of our study, several limitations should be noted. Firstly, the retrospective nature of the analysis may introduce inherent biases related to data collection and patient selection. Secondly, the study was conducted in a single institution, which may limit the generalizability of the findings to other settings or populations. Thirdly, although we employed a variety of machine learning algorithms, the study's performance metrics may still be affected by the relatively small sample size (555 patients), especially when dealing with a rare outcome like postoperative pneumonia. Fourthly, our dataset lacked certain potentially influential variables, such as detailed nutritional status, preoperative functional status, and intraoperative factors, which could further refine the predictive model. Finally, while the eXGBM model showed the best performance, external validation in different cohorts is necessary to confirm its robustness and applicability in broader clinical practice.

Conclusions

This study develops and validates a machine learning-based model to predict postoperative pneumonia in super-aged patients with hip fractures. Utilizing a range of advanced algorithms, the eXGBM model emerges as the most effective, demonstrating superior predictive performance and calibration. The identified key variables, including RDW, blood creatinine, surgery type, and hemoglobin, among others, provide a strong basis for accurate risk stratification. Despite the study's limitations, the model holds promise as a valuable clinical tool to guide preoperative planning and postoperative management, ultimately improving patient outcomes in this vulnerable population. Further external validation is recommended to ensure the model's generalizability and robustness across diverse clinical settings.

Data Sharing Statement

The data can be obtained from the corresponding authors after reasonable request.

Ethics Approval and Consent to Participate

The Medical Research Ethics Board of the Chinese PLA General Hospital approved the study protocol and waived the need for patient consent for the review of medical records and images.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was funded by National Clinical Research Center for Orthopedics, Sports Medicine and Rehabilitation of China, National Key Research and Development Program of China (No. 2022YFC2504300), Beijing Natural Science Foundation (No. 7254420), Hainan Province Clinical Medical Center, and Military Training Injury Prevention and Control Special Project (No. 21XLS38).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Walter N, Szymski D, Kurtz SM, et al. Epidemiology and treatment of proximal femoral fractures in the elderly U.S. population. *Sci Rep*. 2023;13(1):12734. doi:10.1038/s41598-023-40087-8
2. Kawanishi H, Shintaku S, Moriishi M. Vascular access in super-aged patients. *J Vasc Access*. 2015;16(Suppl 10):S22–27. doi:10.5301/jva.5000428
3. Kanda Y, Kakutani K, Sakai Y, et al. Clinical characteristics and surgical outcomes of metastatic spine tumors in the very elderly: a prospective cohort study in a super-aged society. *J Clin Med*. 2023;12(14):4747. doi:10.3390/jcm12144747
4. Chen M, Du Y, Tang W, et al. Risk factors of mortality and second fracture after elderly Hip fracture surgery in Shanghai, China. *J Bone Miner Metab*. 2022;40(6):951–959. doi:10.1007/s00774-022-01358-y
5. Gao YC, Zhang YW, Shi L, et al. What are risk factors of postoperative pneumonia in geriatric individuals after hip fracture surgery: a systematic review and meta-analysis. *Orthop Surg*. 2023;15(1):38–52. doi:10.1111/os.13631
6. Fukuda T, Imai S, Shimoda S, Maruo K, Nakadera M, Horiguchi H. Aspiration pneumonia and anesthesia techniques in Hip fracture surgery in elderly patients: a retrospective cohort study using administrative data. *J Orthop Surg*. 2022;30(1):10225536221078622. doi:10.1177/10225536221078622
7. Han SB, Kim SB, Shin KH. Risk factors for postoperative pneumonia in patients undergoing Hip fracture surgery: a systematic review and meta-analysis. *BMC Musculoskelet Disord*. 2022;23(1):553. doi:10.1186/s12891-022-05497-1
8. Tian Y, Zhu Y, Zhang K, et al. Incidence and risk factors for postoperative pneumonia following surgically treated Hip fracture in geriatric patients: a retrospective cohort study. *J Orthop Surg Res*. 2022;17(1):179. doi:10.1186/s13018-022-03071-y
9. Meng Y, Liu Y, Fu M, Hou Z, Wang Z. Clinical characteristics of elderly Hip fracture patients with chronic cerebrovascular disease and construction of a clinical predictive model for perioperative pneumonia. *Orthop Traumatol Surg Res*. 2024;110(3):103821. doi:10.1016/j.otsr.2024.103821
10. Zhang X, Shen ZL, Duan XZ, et al. Postoperative pneumonia in geriatric patients with a hip fracture: incidence, risk factors and a predictive nomogram. *Geriatr Orthop Surg Rehabil*. 2022;13:21514593221083824. doi:10.1177/21514593221083824
11. Li Y, Dong B. Development and validation of risk prediction nomograms for acute respiratory failure in elderly patients with Hip fracture. *J Orthop Surg Res*. 2023;18(1):899. doi:10.1186/s13018-023-04395-z
12. Cheng X, Liu Y, Wang W, et al. Preoperative risk factor analysis and dynamic online nomogram development for early infections following primary hip arthroplasty in geriatric patients with hip fracture. *Clin Interv Aging*. 2022;17:1873–1883. doi:10.2147/CIA.S392393
13. Wang B, Li Y, Tian Y, Ju C, Xu X, Pei S. Novel pneumonia score based on a machine learning model for predicting mortality in pneumonia patients on admission to the intensive care unit. *Respir Med*. 2023;217:107363. doi:10.1016/j.rmed.2023.107363

14. Tian CW, Chen XX, Shi L, et al. Machine learning applications for the prediction of extended length of stay in geriatric Hip fracture patients. *World J Orthop*. 2023;14(10):741–754. doi:10.5312/wjo.v14.i10.741
15. Rajula HSR, Verlato G, Manchia M, Antonucci N, Fanos V. Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicina*. 2020;56(9). doi:10.3390/medicina56090455
16. Sulaiman S, Kawsara A, El Sabbagh A, et al. Machine learning vs. conventional methods for prediction of 30-day readmission following percutaneous mitral edge-to-edge repair. *Cardiovasc Revasc Med*. 2023;56:18–24. doi:10.1016/j.carrev.2023.05.013
17. Guo J, He Q, Peng C, et al. Machine learning algorithms to predict risk of postoperative pneumonia in elderly with Hip fracture. *J Orthop Surg Res*. 2023;18(1):571. doi:10.1186/s13018-023-04049-0
18. Dai A, Liu H, Shen P, et al. Incorporating preoperative frailty to assist in early prediction of postoperative pneumonia in elderly patients with Hip fractures: an externally validated online interpretable machine learning model. *BMC Geriatr*. 2024;24(1):472. doi:10.1186/s12877-024-05050-w
19. Mathew G, Agha R, Albrecht J, et al. STROCSS 2021: strengthening the reporting of cohort, cross-sectional and case-control studies in surgery. *Int J Surg*. 2021;96:106165. doi:10.1016/j.ijsu.2021.106165
20. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD). *Ann Intern Med*. 2015;162(10):735–736. doi:10.7326/L15-5093-2
21. Shi X, Cui Y, Wang S, Pan Y, Wang B, Lei M. Development and validation of a web-based artificial intelligence prediction model to assess massive intraoperative blood loss for metastatic spinal disease using machine learning techniques. *Spine J*. 2024;24(1):146–160. doi:10.1016/j.spinee.2023.09.001
22. Lei M, Feng T, Chen M, et al. Establishment and validation of an artificial intelligence web application for predicting postoperative in-hospital mortality in patients with Hip fracture: a national cohort study of 52 707 cases. *Int J Surg*. 2024;110(8):4876–4892. doi:10.1097/JS9.0000000000001599
23. Liu X, Guo L, Wang H, Guo J, Yang S, Duan L. Research on imbalance machine learning methods for MR[Formula: see text]WI soft tissue sarcoma data. *BMC Med Imaging*. 2022;22(1):149. doi:10.1186/s12880-022-00876-5
24. Zhu C, Xu Z, Gu Y, et al. Prediction of post-stroke urinary tract infection risk in immobile patients using machine learning: an observational cohort study. *J Hosp Infect*. 2022;122:96–107. doi:10.1016/j.jhin.2022.01.002
25. Sun B, Lei M, Wang L, et al. Prediction of sepsis among patients with major trauma using artificial intelligence: a multicenter validated cohort study. *Int J Surg*. 2024.
26. Ahn J, Chang JS, Kim JW. Postoperative pneumonia and aspiration pneumonia following elderly hip fractures. *J Nutr Health Aging*. 2022;26(7):732–738. doi:10.1007/s12603-022-1821-9
27. Yao W, Wang W, Tang W, Lv Q, Ding W. Neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and systemic immune inflammation index (SII) to predict postoperative pneumonia in elderly Hip fracture patients. *J Orthop Surg Res*. 2023;18(1):673. doi:10.1186/s13018-023-04157-x
28. Zhang D, Zhang Y, Yang S, Sun L, Zhang N, Huang S. Relationship between preoperative red blood cell distribution width and postoperative pneumonia in elderly patients with Hip fracture: a retrospective cohort study. *J Orthop Surg Res*. 2023;18(1):253. doi:10.1186/s13018-023-03732-6
29. Tian Y, Zhu Y, Zhang K, Tian M, Qin S, Li X. Relationship between preoperative hypoalbuminemia and postoperative pneumonia following geriatric hip fracture surgery: a propensity-score matched and conditional logistic regression analysis. *Clin Interv Aging*. 2022;17:495–503. doi:10.2147/CIA.S352736
30. Kraaijkamp JJM, Persoon A, Aurelian S, et al. eHealth in geriatric rehabilitation: an international survey of the experiences and needs of healthcare professionals. *J Clin Med*. 2023;12(13):4504. doi:10.3390/jcm12134504