



Case Control Study

Machine learning applications for the prediction of extended length of stay in geriatric hip fracture patients

Chu-Wei Tian, Xiang-Xu Chen, Liu Shi, Huan-Yi Zhu, Guang-Chun Dai, Hui Chen, Yun-Feng Rui

Specialty type: Orthopedics

Provenance and peer review:

Unsolicited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): 0

Grade B (Very good): B, B, B

Grade C (Good): C

Grade D (Fair): 0

Grade E (Poor): 0

P-Reviewer: Lee KS, South Korea;
Mijwil MM, Iraq

Received: July 22, 2023

Peer-review started: July 22, 2023

First decision: September 4, 2023

Revised: September 8, 2023

Accepted: September 28, 2023

Article in press: September 28, 2023

Published online: October 18, 2023



Chu-Wei Tian, Xiang-Xu Chen, Liu Shi, Huan-Yi Zhu, Guang-Chun Dai, Hui Chen, Yun-Feng Rui,
Department of Orthopaedics, Zhongda Hospital, School of Medicine, Southeast University,
Nanjing 210009, Jiangsu Province, China

Chu-Wei Tian, Xiang-Xu Chen, Liu Shi, Huan-Yi Zhu, Guang-Chun Dai, Hui Chen, Yun-Feng Rui,
Multidisciplinary Team for Geriatric Hip Fracture Comprehensive Management, Zhongda
Hospital, School of Medicine, Southeast University, Nanjing 210009, Jiangsu Province, China

Chu-Wei Tian, Xiang-Xu Chen, Liu Shi, Huan-Yi Zhu, Guang-Chun Dai, Hui Chen, Yun-Feng Rui,
Trauma Center, Zhongda Hospital, School of Medicine, Southeast University, Nanjing 210009,
Jiangsu Province, China

Corresponding author: Yun-Feng Rui, MD, PhD, Chief Doctor, Deputy Director, Professor,
Surgeon, Department of Orthopaedics, Zhongda Hospital, School of Medicine, Southeast
University, No. 87 Dingjiaqiao, Nanjing 210009, Jiangsu Province, China.
ruiyunfeng@126.com

Abstract

BACKGROUND

Geriatric hip fractures are one of the most common fractures in elderly individuals, and prolonged hospital stays increase the risk of death and complications. Machine learning (ML) has become prevalent in clinical data processing and predictive models. This study aims to develop ML models for predicting extended length of stay (eLOS) among geriatric patients with hip fractures and to identify the associated risk factors.

AIM

To develop ML models for predicting the eLOS among geriatric patients with hip fractures, identify associated risk factors, and compare the performance of each model.

METHODS

A retrospective study was conducted at a single orthopaedic trauma centre, enrolling all patients who underwent hip fracture surgery between January 2018 and December 2022. The study collected various patient characteristics, encompassing demographic data, general health status, injury-related data, laboratory examinations, surgery-related data, and length of stay. Features that exhibited significant differences in univariate analysis were integrated into the ML model

establishment and subsequently cross-verified. The study compared the performance of the ML models and determined the risk factors for eLOS.

RESULTS

The study included 763 patients, with 380 experiencing eLOS. Among the models, the decision tree, random forest, and extreme Gradient Boosting models demonstrated the most robust performance. Notably, the artificial neural network model also exhibited impressive results. After cross-validation, the support vector machine and logistic regression models demonstrated superior performance. Predictors for eLOS included delayed surgery, D-dimer level, American Society of Anaesthesiologists (ASA) classification, type of surgery, and sex.

CONCLUSION

ML proved to be highly accurate in predicting the eLOS for geriatric patients with hip fractures. The identified key risk factors were delayed surgery, D-dimer level, ASA classification, type of surgery, and sex. This valuable information can aid clinicians in allocating resources more efficiently to meet patient demand effectively.

Key Words: Machine learning; Extended length of stay; Hip fracture; Enhanced recovery after surgery; Risk factors

©The Author(s) 2023. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Traditional models have been built to identify risk factors for extended length of stay (eLOS), offering new insights for optimizing treatment for hip fracture patients under the enhanced recovery after surgery concept. However, these traditional statistical methods suffer from poor performance and lack of features. Machine learning (ML) is a scientific discipline focused on teaching computers to learn from data, showing superior predictive performance compared to traditional methods. This study aims to develop ML models for predicting eLOS among geriatric patients with hip fractures, identify associated risk factors, and compare the performance of each model.

Citation: Tian CW, Chen XX, Shi L, Zhu HY, Dai GC, Chen H, Rui YF. Machine learning applications for the prediction of extended length of stay in geriatric hip fracture patients. *World J Orthop* 2023; 14(10): 741-754

URL: <https://www.wjgnet.com/2218-5836/full/v14/i10/741.htm>

DOI: <https://dx.doi.org/10.5312/wjo.v14.i10.741>

INTRODUCTION

Hip fractures have become more prevalent as the global geriatric population increases[1]. They are associated with higher incidence, mortality, and disability, significantly impacting the quality of life of affected individuals[2,3]. Prolonged length of stay (LOS) not only places a financial burden on patients but also elevates the risk of mortality and complications[4]. Enhanced recovery after surgery (ERAS) refers to the integration of perioperative concepts using evidence-based medicine tools to reduce surgical stress and complications, shorten hospital stays, lower financial costs, and hasten postoperative recovery[5-7]. Based on this concept, Andrew *et al*[8] developed a logistic regression (LR) model to identify risk factors for extended length of stay (eLOS), offering new insights for optimizing treatment for hip fracture patients. However, traditional statistical methods suffer from poor performance and lack of features.

Machine learning (ML) is a scientific discipline focused on teaching computers to learn from data[9]. In recent times, ML has shown superior predictive performance compared to traditional methods and has found extensive application in clinical data processing and predictive modelling[10,11]. Mijwil and Aggarwal[12] enhanced the ML based estimation method for detecting acute appendicitis in individuals, achieving high accuracy. In the context of hip fractures among geriatric individuals, Oosterhoff *et al*[13] and Shtar *et al*[14] established ML models to predict prognosis and mortality, enhancing clinician decision-making ability. With the establishment of a clinical database for elderly hip fracture patients and advancements in ML algorithms, predicting the eLOS for this patient group through ML has become feasible.

This study aims to develop ML models for predicting the eLOS among geriatric patients with hip fractures, identify associated risk factors, and compare the performance of each model.

MATERIALS AND METHODS

Study design, setting and population

A retrospective study was conducted at a single orthopaedic trauma centre between January 2018 and December 2022. The study employed specific inclusion and exclusion criteria as follows inclusion criteria: (1) Age older than 60 years at the time of injury; (2) confirmed diagnosis of hip fracture; and (3) hospitalization at our centre. The study employed specific inclusion and exclusion criteria as follows exclusion criteria: (1) Admission with multiple fractures, pathological

fractures, or fractures around the prosthesis; (2) receipt of conservative treatment due to severe comorbidities; and (3) presence of missing data.

The enrolled patients had a median hospital stay of 9.5 d. Based on this median LOS, the patients were retrospectively divided into two groups: None LOS (LOS \leq 9.5 d, $n = 383$, 50.2%) and eLOS (LOS $>$ 9.5 d, $n = 380$, 49.8%).

Perioperative treatment and surgical procedure

All patients underwent comprehensive perioperative assessments to ensure standardized diagnosis and treatment. Our centre boasts a multidisciplinary team comprising geriatricians, anaesthetists, and intensive care unit (ICU) doctors who collaborate to review the perioperative care of patients with comorbidities[15]. Upon admission, a rapid preoperative risk assessment is performed, considering the patient's physical condition and specific needs. Subsequently, an individualized treatment plan is promptly devised. For patients requiring surgical intervention, a consistent surgical team oversees all procedures. Patients with femoral neck fractures receive surgical treatments such as total hip arthroplasty (THA), hemiarthroplasty of the hip, or internal fixation. Those with intertrochanteric fractures undergo surgical treatment, such as internal fixation. Throughout the surgery, the patient's temperature, blood volume, and haemodynamics are meticulously managed with the collaborative efforts of anaesthesia and surgical nurses.

Data collection

Data for the study were retrospectively gathered from electronic patient records at the institution. Demographic data encompassed sex, age, body mass index, general health status categorized by the American Society of Anaesthesiologists (ASA) classification, history of smoking, oral anticoagulant use, and comorbidities[16]. Injury-related data included fracture type, time from injury to admission, and the day of admission. Surgery-related data consisted of the type of surgery, anaesthesia used, ICU transfer, time to surgery, duration of surgery, intraoperative blood loss, and transfusion. Laboratory examinations conducted at admission and after surgery were also collected. Age was stratified into 60-85 and $>$ 85 age groups; ASA classification was grouped as I-II and III-IV; admission day was grouped into Monday to Thursday and Friday to Sunday; injury time was stratified into \leq 24 and $>$ 24 h; and delayed surgery was defined as an operation performed more than 48 h after admission. Laboratory examinations were stratified according to normal values.

ML and statistical analysis

In the study, normally distributed data are presented as the means and standard deviations. Nonnormally distributed variables were expressed as medians along with their interquartile ranges. Categorical variables are represented as counts and percentages. To analyse the overall data, continuous variables were subjected to Student's t test or the Mann-Whitney U test, depending on their distribution. Categorical variables were analysed using the chi-square test, as appropriate. Variables showing significant differences in the univariate analysis were selected and included in the establishment of the ML model.

The predictive eLOS ML model was established according to the selected features, including basic algorithms for LR, decision tree (DT), random forest (RF), support vector machine (SVM), naïve Bayes, K-nearest neighbour, eXtreme Gradient Boosting (XGBoost) and artificial neural network (ANN) models. Next, we used a Shapley Additive Interpretation (SHAP) summary plot to determine the relationship between the eLOS and its main predictors in each model. Each ML model was integrated to ascertain feature importance. Then, the original data were split into a training set and a test set (training: test = 7:3), and 10-fold cross-validation was carried out. The accuracy score, precision score, recall score, F1 score, receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) were used to evaluate the performance of the ML model of the original data and cross-validation. All statistical analyses were conducted using Python (version 3.8.2, Python Software Foundation, <https://www.python.org>) and the sklearn package (version 0.24.1). A 2-sided P value $<$ 0.05 was considered significant. The flow diagram of the research process is shown in Figure 1.

RESULTS

Population and patient characteristics

Overall, 763 patients were enrolled in the final analysis; patients were divided into none LOS ($n = 383$) and eLOS ($n = 380$) groups based on LOS. The characteristics of the two groups are compared in Table 1. Univariate analysis showed that there were significant differences in sex, fracture type, ASA classification, admission day, injury to admission time, hypertension, diabetes, cerebral infarction, deep venous thrombosis (DVT), delayed surgery, THA, reduction and fixation, aspartate aminotransferase and D-dimer levels and aortic velocity at admission between the two groups ($P <$ 0.05).

Establishment and evaluation of the ML model in the original data

We used 8 ML models to evaluate the predictors of LOS in the original data. Figure 2 shows the ROC curve, and Table 2 shows the performance indicators of each model. The tree models, including the DT (accuracy = 0.924, AUC = 0.988), RF (accuracy = 0.924, AUC = 0.985) and XGBoost (accuracy = 0.912, AUC = 0.976) models, showed stronger performance among the models. In addition, the performance of the ANN (accuracy = 0.886, AUC = 0.963) model was impressive.

Table 1 Baseline data comparison. Counts (%) unless otherwise specified

Features	Non-eLOS (n = 383)	eLOS (n = 380)	Statistic (t/χ^2)	P value
Sex			5.418	0.020
Male	107 (27.9)	136 (35.8)		
Female	276 (72.1)	244 (64.2)		
Age (yr)			0.096	0.757
60-85	255 (66.6)	257 (67.6)		
> 85	128 (33.4)	123 (32.4)		
BMI (kg/m ²)			2.578	0.461
< 18.5	48 (12.5)	44 (11.6)		
18.5-24.9	220 (57.4)	239 (62.9)		
25.0-29.9	99 (25.8)	82 (21.6)		
> 30.0	16 (4.2)	15 (3.9)		
Fracture type			20.516	P < 0.001
Femoral neck fracture	165 (43.1)	226 (59.5)		
Intertrochanteric fracture	218 (56.9)	154 (40.5)		
ASA classification			20.313	P < 0.001
I-II	197 (51.4)	134 (35.3)		
III-IV	186 (48.6)	246 (64.7)		
Admission time			5.340	0.021
Monday to Thursday	248 (64.8)	215 (56.6)		
Friday to Sunday	135 (35.2)	165 (43.2)		
Time from injury to admission			5.513	0.019
≤ 24 h	322 (84.1)	294 (77.4)		
> 24 h	61 (15.9)	86 (22.6)		
Comorbidities	3.99 2.22	4.19 2.25	1.221	0.223
Hypertension	207 (54.0)	245 (64.5)	8.588	0.003
Diabetes	85 (22.2)	112 (29.5)	5.279	0.022
ACS	59 (15.4)	66 (17.4)	0.537	0.464
Cerebral infarction	120 (31.3)	150 (39.5)	5.531	0.019
AKI	8 (2.1)	17 (4.5)	3.423	0.064
DVT	66 (17.2)	91 (23.9)	5.263	0.022
History of smoking	14 (3.4)	17 (4.5)	0.328	0.567
Oral anticoagulant use	71 (18.5)	71 (18.7)	0.003	0.959
History of fracture	77 (20.1)	83 (21.8)	0.348	0.556
Hip fracture	21 (5.5)	29 (7.6)	1.438	0.230
Lumbar fracture	25 (6.5)	24 (6.3)	0.014	0.905
Delayed surgery	153 (39.9)	267 (70.3)	70.842	P < 0.001
Type of surgery				
THA	52 (13.6)	95 (25.0)	16.002	P < 0.001
HHA	106 (27.7)	127 (33.4)	2.968	0.085
Reduction and fixation	225 (58.7)	158 (41.6)	22.488	P < 0.001
Duration of surgery	96.62 ± 27.83	118.43 ± 24.32	1.743	0.082

Intraoperative blood loss	164.54 ± 85.26	162.33 ± 104.20	-0.321	0.748
Blood transfusion	127 (33.2)	132 (34.7)	0.212	0.645
ICU transfer	103 (26.9)	111 (29.2)	0.508	0.476
Type of anaesthesia			0.769	0.380
General	336 (87.7)	341 (89.7)		
Regional	47 (12.3)	39 (10.3)		
Heart rate at admission (60-100)	339 (88.5)	331 (87.1)	0.353	0.553
Laboratory examination at admission				
RBC (≥ 4.3)	86 (22.5)	87 (22.9)	0.021	0.884
WBC (3.5-9.5)	208 (54.3)	217 (57.1)	0.605	0.437
Hb (≥ 110)	251 (65.5)	268 (70.5)	2.184	0.139
PLT (125-350)	321 (83.8)	306 (80.5)	1.406	0.236
N (40-75)	65 (17.0)	81 (21.3)	2.327	0.127
HCT (≥ 40)	82 (21.4)	65 (17.1)	2.272	0.132
K (3.5-5.1)	260 (67.9)	265 (69.7)	0.305	0.581
Ca (≥ 2.1)	257 (67.1)	260 (68.4)	0.152	0.697
Na (137-145)	248 (64.8)	229 (60.3)	1.640	0.200
ALB (30-40)	233 (60.8)	243 (63.9)	0.787	0.375
ALT (9-50)	347 (90.5)	342 (90.0)	0.079	0.779
AST (15-40)	338 (88.3)	316 (83.2)	4.040	0.044
LDH (120-246)	188 (49.1)	179 (47.1)	3.000	0.584
BUN (3.6-9.5)	304 (79.4)	302 (79.5)	0.001	0.973
Cr (58-110)	205 (53.5)	221 (58.2)	1.660	0.198
PT (9.4-12.5)	77 (20.1)	93 (24.5)	2.103	0.147
APTT (25.1-36.5)	326 (85.1)	331 (87.1)	0.630	0.427
INR (0.8-1.2)	311 (81.2)	304 (80.0)	0.176	0.675
FIB (2.38-4.98)	344 (89.8)	344 (90.5)	0.108	0.742
D-dimer (≤ 6500)	218 (56.9)	165 (43.4)	13.902	<i>P</i> < 0.001
Cardiac colour ultrasound at admission				
AV (≥ 1.0)	367 (95.8)	347 (91.3)	6.446	0.011
EF (≥ 70)	208 (54.3)	217 (57.1)	0.605	0.437
Laboratory examination after surgery				
Hb (≥ 110)	87 (22.7)	92 (24.2)	0.237	0.626
ALB (30-40)	310 (80.9)	303 (79.7)	0.175	0.676
Cr (58-110)	222 (58.0)	232 (61.1)	0.755	0.385

BMI: Body mass index; ASA: American society of anaesthesiologists; ACS: Acute coronary syndromes; AKI: Acute kidney injury; DVT: Deep venous thrombosis; THA: Total hip arthroplasty; HHA: Hip hemiarthroplasty; ICU: Intensive care unit; RBC: Red blood cell; WBC: White blood cell; Hb: Hemoglobin; PLT: Platelets; N: Neutrophil granulocyte; HCT: Hematocrit; K: Kalium; Ca: Calcium; Na: Natrium; ALB: Albumin; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; LDH: Lactate dehydrogenase; BUN: Blood urea nitrogen; Cr: Creatinine; PT: Prothrombin time; APTT: Activated partial thromboplastin time; INR: International normalized ratio; FIB: Fibrinogen; AV: Aortic velocity; EF: Ejection fraction.

Table 2 Evaluation of machine learning models in the original data

Model name	Accuracy	Precision	Recall	F1 score	AUC
LR	0.680	0.680	0.676	0.678	0.747
DT	0.924	0.994	0.853	0.918	0.988
RF	0.924	0.940	0.905	0.922	0.985
SVM	0.651	0.636	0.703	0.667	0.739
NB	0.657	0.676	0.598	0.634	0.709
KNN	0.747	0.748	0.742	0.745	0.828
XGB	0.912	0.941	0.879	0.901	0.976
ANN	0.886	0.899	0.868	0.883	0.963

AUC: Area under curve; LR: Logistic regression; DT: Decision tree; RF: Random forest; SVM: Support vector machine; NB: Naïve bayes; KNN: K-nearest Neighbour; XGB: eXtreme Gradient Boosting; ANN: Artificial neural network.

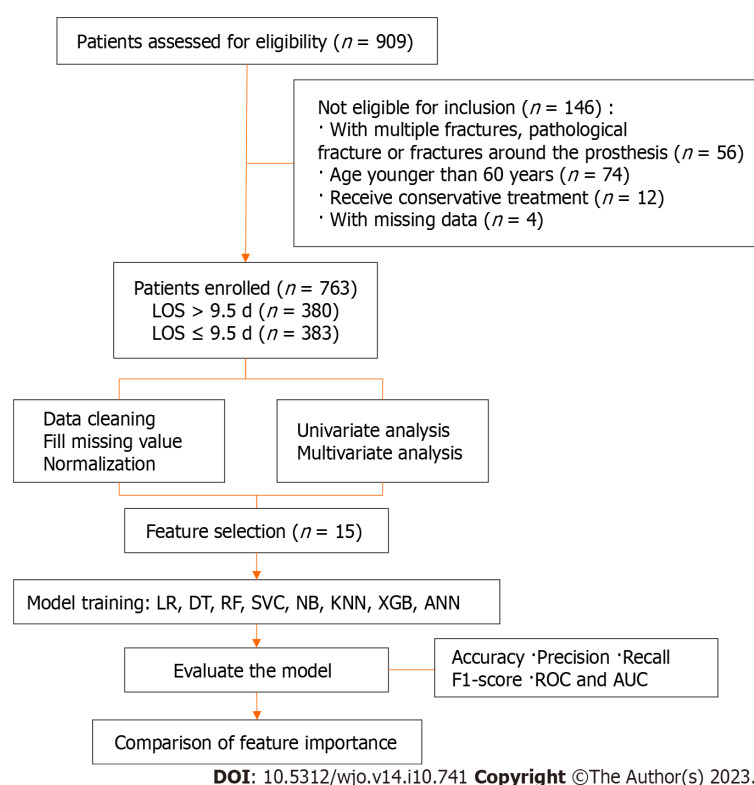


Figure 1 The flow diagram of the research process. LOS: Length of stay; LR: Logistic regression; DT: Decision tree; RF: Random forest; SVC: Support vector classifier; NB: Naïve bayes; KNN: K-nearest neighbour; XGB: eXtreme Gradient Boosting; ANN: Artificial neural network; ROC: Receiver operating characteristic; AUC: area under the receiver operating characteristic curve.

Feature importance

Figure 3 shows the results of the SHAP analysis. We can intuitively understand the importance of features in each model and the direction of their association with the eLOS. Then, we summarized the importance of the features output by each model. Figure 4 shows the feature weights in descending order; delayed surgery was the most important feature in eLOS prediction. The other most important features that influenced the prediction of the eLOS were D-dimer level, ASA classification, type of surgery and sex.

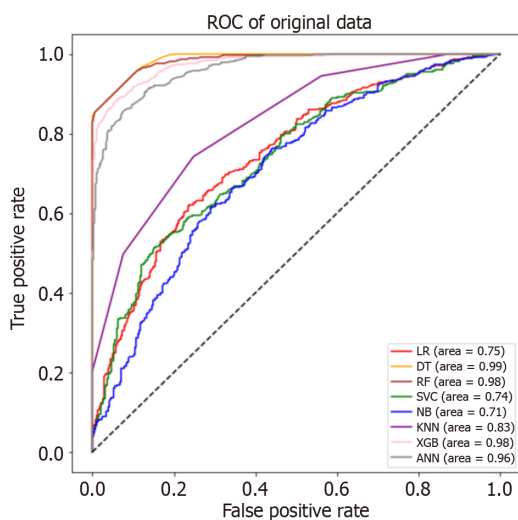
Evaluation of ML models after 10-fold cross-validation

We split the original data into training and test sets (training: test = 7:3) and carried out 10-fold cross-validation. Figure 5 shows the ROC curve, and Table 3 indicates the performance indicators of each model after cross-validation. We found that the performance of the model decreased after cross-validation compared to the original data. The best results were found by using the SVM (accuracy = 0.664, AUC = 0.712) model. Furthermore, the LR (accuracy = 0.650, AUC = 0.650)

Table 3 Evaluation of machine learning models after 10-fold cross-validation

Model name	Accuracy	Precision	Recall	F1 score	AUC
LR	0.650	0.643	0.668	0.655	0.650
DT	0.606	0.616	0.552	0.583	0.605
RF	0.619	0.618	0.613	0.616	0.619
SVM	0.664	0.656	0.659	0.658	0.712
NB	0.644	0.657	0.595	0.624	0.643
KNN	0.617	0.611	0.639	0.625	0.617
XGB	0.630	0.632	0.616	0.624	0.630
ANN	0.606	0.596	0.645	0.619	0.606

AUC: Area under curve; LR: Logistic regression; DT: Decision tree; RF: Random forest; SVM: Support vector machine; NB: Naïve bayes; KNN: K-nearest Neighbour; XGB: eXtreme Gradient Boosting; ANN: Artificial neural network.



DOI: 10.5312/wjo.v14.i10.741 Copyright ©The Author(s) 2023.

Figure 2 Receiver operating characteristic curve for machine learning models in the original data. ROC: Receiver operating characteristic; LR: Logistic regression; DT: Decision tree; RF: Random forest; SVC: Support vector classifier; NB: Naïve Bayes; KNN: K-nearest Neighbour; XGB: eXtreme Gradient Boosting; ANN: Artificial neural network.

model showed good performance after cross-validation.

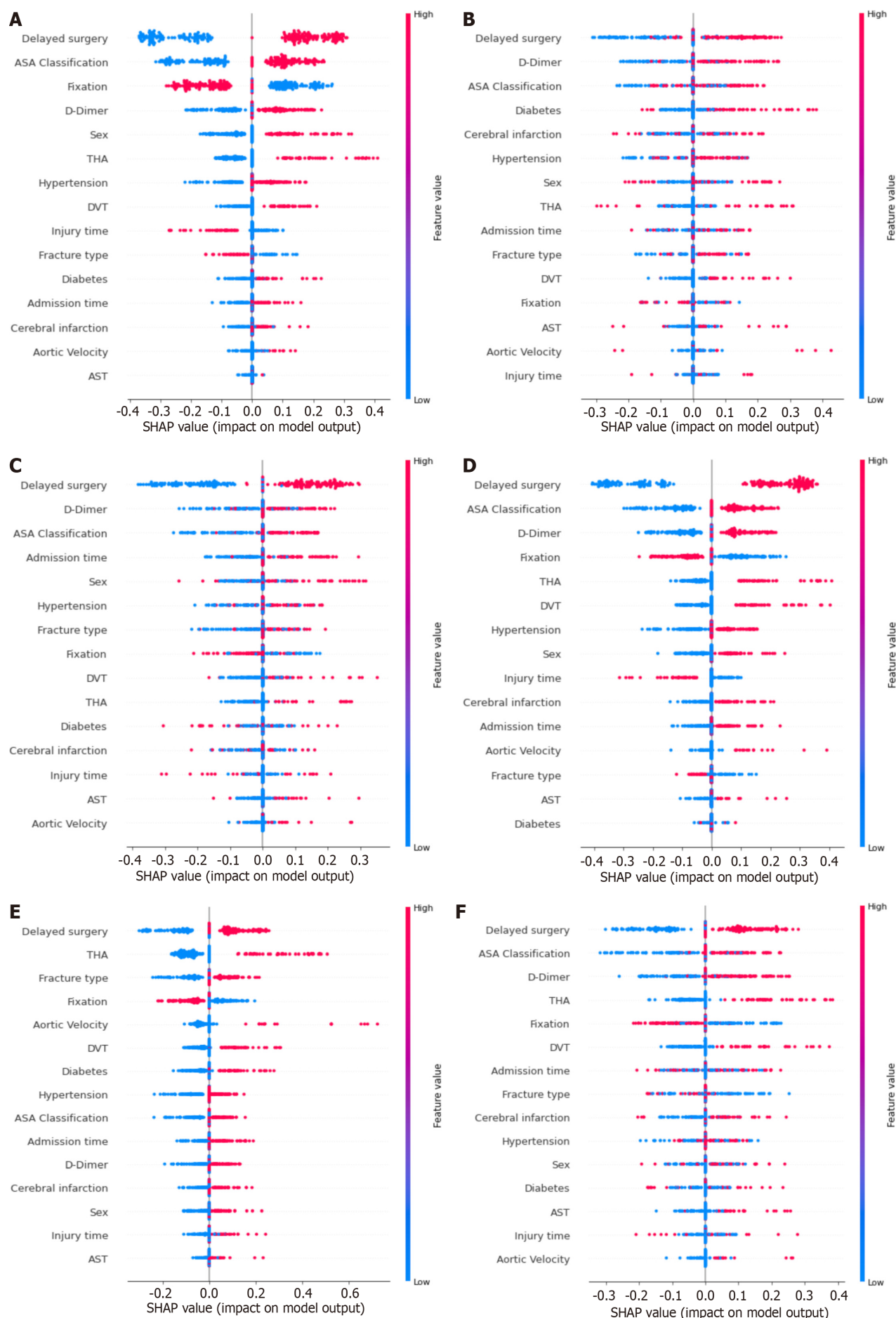
DISCUSSION

This study aimed to develop ML models for predicting eLOS in geriatric patients with hip fractures and to identify associated risk factors. Additionally, we assessed and compared the performance of each ML model. The DT and ANN models demonstrated the best performance with the original data. After cross-validation, the SVM and LR models also performed well.

Risk factors for eLOS among geriatric hip fracture patients

Although methods such as multidisciplinary management and ERAS have been shown to decrease LOS and lower inpatient hospitalization costs, geriatric hip fracture patients continue to be disproportionately large resource consumers [8,17,18]. Long-term LOS not only results in inefficient use of medical resources but also increases the risk of complications among hip fracture patients [19]. In addition, the length of hospital stay varies greatly among healthcare systems [17,18,20]. The identification of risk factors associated with the eLOS may be helpful in cost forecasting and patient management [21]. In our study, we used the median LOS to group the data.

In previous studies, delayed surgery was considered the key factor that affected LOS. A retrospective study by Pincus *et al* [22] showed that patients who underwent surgery 1 day after admission had a longer postoperative stay. Hecht *et al* [23] developed a predictive model for LOS that showed that the prevention of reduced time to surgery was a significant



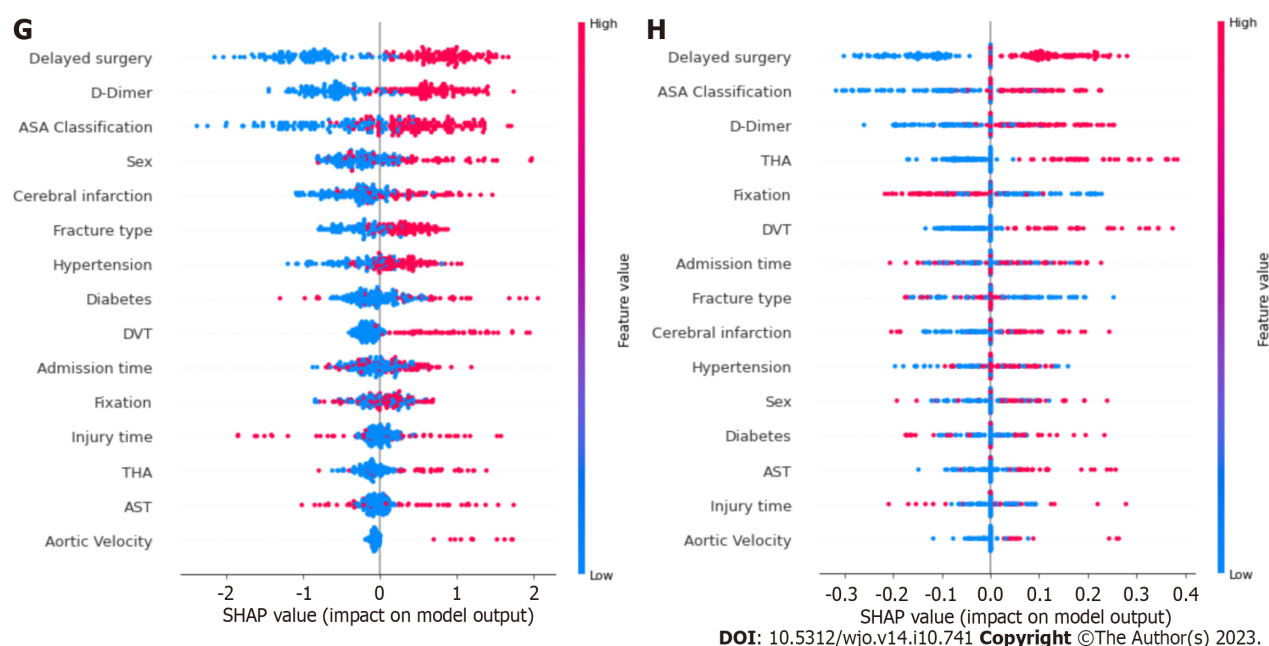


Figure 3 Shapley additive explanations summary plots of each model. A: Logistic regression; B: Decision tree; C: Random forest; D: Support vector classifier; E: Naïve bayes; F: K-nearest neighbour; G: eXtreme Gradient Boosting; H: Artificial neural network. ASA: American society of anaesthesiologists; THA: Total hip arthroplasty; DVT: Deep venous thrombosis; AST: Aspartate aminotransferase.

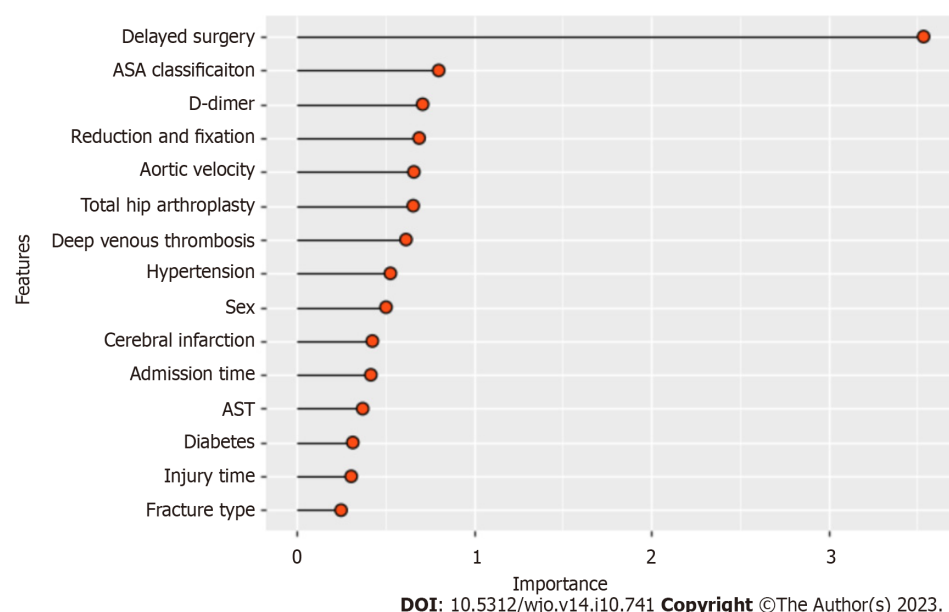
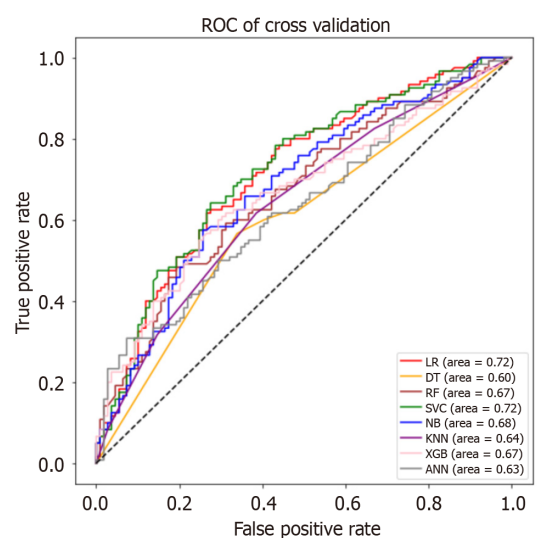


Figure 4 Comprehensive importance of features. ASA: American society of anaesthesiologists; AST: Aspartate aminotransferase.

predictor of reduced LOS. In addition, the authors suggested that ASA classification was a stronger predictor of LOS than the Charlson comorbidity index. Similar findings were shown by Kristan *et al*[24], who found that the eLOS was associated with delayed surgery, ASA classification, anticoagulant therapy and surgery type. Furthermore, a higher D-dimer level, as a possible predictor of DVT, suggested that patients were at higher risk of thrombosis, which was also the cause of eLOS[25]. Our study achieved similar results to those described above. This result suggests that clinicians and multidisciplinary teams should continue to explore possible interventions to shorten LOS among hip fracture patients.

At the same time, our model identified male sex and fracture type as predictors of the eLOS. This is consistent with findings from the study by Garcia *et al*[26]. The results of their study showed that while the majority of hip fracture patients were female, male patients appeared to have a longer hospital stay. A meta-analysis by Haentjens *et al*[27] showed that male hip fracture patients had a higher risk of death, which appeared to be associated with more severe osteoporosis and a higher comorbidity burden among male patients with hip fractures[28]. In addition, patients with joint replacement surgery had higher functional requirements, which was why such patients took longer to stay and recover [29,30]. However, some studies mentioned the influence of age and comorbidities on the eLOS, but these factors were not



DOI: 10.5312/wjo.v14.i10.741 Copyright ©The Author(s) 2023.

Figure 5 Receiver operating characteristic curve for machine learning models after cross-validation. ROC: Receiver operating characteristic; LR: Logistic regression; DT: Decision tree; RF: Random forest; SVC: Support vector classifier; NB: Naïve bayes; KNN: K-nearest Neighbour; XGB: eXtreme Gradient Boosting; ANN: Artificial neural network.

identified by our model[24,27,30]. This might be due to changes in the management of hip fracture patients and the popularity of ERAS, which have allowed an increasing number of older patients with comorbidities to receive timely surgical treatment and rehabilitation guidance.

Predictive performance of ML models for eLOS

With the development of science and technology, ML is also being used in the field of medicine to improve patient outcomes and diagnostic accuracy[31]. Recently, ML has been widely used in the diagnosis, classification, identification and prognosis of hip fracture patients[32-34]. Promising results were obtained by Forssten *et al*[35], who used ML to predict 1-year mortality after hip fracture surgery, and by Galassi *et al*[36], who used ML to assess hip fracture risk. With the establishment of clinical databases, ML models will have better practical value in the future.

Most of the previous studies on the construction of prediction models were based on regression algorithms[9]. For binary clinical decision data, the tree model has a natural advantage[37]. The RF algorithm and the XGBoost integration algorithm also show similar results[38]. Based on gradient-boosted DTs, the XGBoost algorithm applied a second-order Taylor expansion to calculate the loss function and performed well in both computational speed and predictive precision [39]. Previous studies have also demonstrated this point. Hou *et al*[40] used XGBoost to predict 30-d mortality for the medical information mart for intensive care III patients with sepsis-3 and obtained high accuracy. Noh *et al*[41] similarly achieved good accuracy in identifying the optimal features of gait parameters to predict fall risk among older adults by XGBoost. In the original data set of our research, the performance of the tree model was far superior to that of other models, which might be the result of overfitting. Through cross-verification, we found that the performance evaluation index of the tree model declined the most. With the continuous expansion of the sample size, the performance of the tree model would also be improved. However, the performance of the traditional binary classification algorithm was stable. In our study, the SVM and LR models had the best performance after cross-validation.

Recently, ANNs have become a new hotspot in ML development. An ANN is a kind of ML algorithm inspired by biological neural networks[10]. Figure 6 shows the computational flow of the ANN in our study (hidden layer size = 15, 10, 10). The ANN contains nodes that communicate with other nodes *via* connections. Chen *et al*[42] showed that the ANN was more accurate than Cox regression in predicting mortality after hip fracture surgery. Using an ANN model can enable more appropriate and accurate processing of inputs that are incomplete or inputs that introduce noise[43,44]. Moreover, the results of Klemm *et al*[45] demonstrated its good application in binary data. In our research, the ANN model showed performance second only to that of the tree model in terms of original data and had considerable accuracy after cross-verification.

Strengths and limitations

This study conducted a novel experiment to develop and compare ML models for predicting eLOS among patients with hip fractures. Subsequently, risk factors for the eLOS were identified. The findings revealed that ML models outperformed traditional statistical methods in terms of accuracy. This provides clinicians with a valuable tool to efficiently identify populations at high risk of eLOS. By doing so, the diagnosis and treatment process can be optimized to reduce the LOS and allocate medical resources effectively, aligning with the ERAS concept.

However, there are several limitations to consider in our study. First, it was a single-centre study, and the length of hospital stay might vary significantly across different healthcare systems. Moreover, the high proportion of patients with ASA III-IV in our hospital indicates a higher prevalence of severe comorbidities and advanced disease compared to those treated in the community, leading to potential selection bias. Second, since this study aimed to establish ML models, the

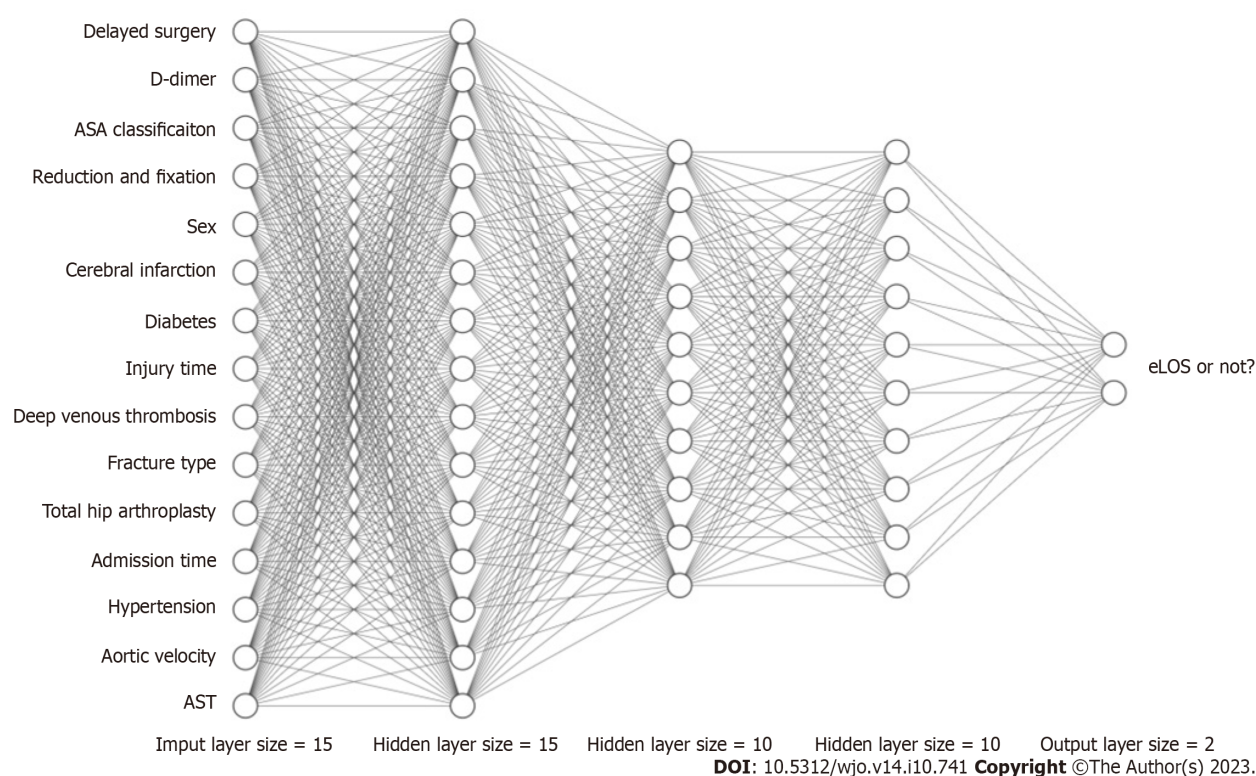


Figure 6 Schematic diagram of an artificial neural network. ASA: American society of anaesthesiologists; eLOS: Extended length of stay; AST: Aspartate aminotransferase.

sample size might be relatively small, resulting in some ML models being prone to overfitting. As a result, the findings of this study should be further validated and made applicable to a broader population through multicentre and large-sample studies.

CONCLUSION

In conclusion, we have effectively developed a highly accurate ML model for eLOS prediction in hip fracture patients. Notably, delayed surgery, elevated D-dimer levels, ASA classification, surgical type, and sex were significantly associated with the eLOS. By applying ML in clinical practice, we can optimize the diagnosis and treatment of elderly hip fracture patients, guide clinicians in decision-making, and allocate medical resources more efficiently.

ARTICLE HIGHLIGHTS

Research background

Geriatric hip fractures are a frequent occurrence and can lead to increased risks of complications and mortality during prolonged hospital stays. This study focuses on utilizing machine learning (ML) to create predictive models aimed at forecasting extended length of stay (eLOS) in elderly patients with hip fractures.

Research motivation

This research endeavor seeks to construct ML models to forecast eLOS in geriatric patients afflicted with hip fractures. Additionally, the study aims to discern the pertinent risk factors contributing to eLOS and conduct a comparative assessment of the performance of each developed model.

Research objectives

This research endeavors to construct ML models for the purpose of forecasting eLOS in geriatric patients who have suffered hip fractures. Furthermore, it seeks to discern the pertinent risk factors associated with this outcome and conduct a comparative analysis of the model performances. We have successfully formulated a highly precise ML model for the prediction of eLOS in patients with hip fractures. Significantly, factors such as delayed surgical intervention, elevated D-dimer levels, American Society of Anaesthesiologists (ASA) classification, surgical procedure type, and gender exhibited notable associations with eLOS. The integration of ML into clinical settings holds the potential to enhance the diagnostic

and therapeutic processes for elderly hip fracture patients, assist clinicians in informed decision-making, and optimize the allocation of healthcare resources.

Research methods

A retrospective investigation was carried out at a sole orthopaedic trauma center, encompassing all individuals who underwent surgery for hip fractures from January 2018 to December 2022. This study compiled a comprehensive array of patient characteristics, encompassing demographics, general health status, injury-related information, laboratory results, surgical data, and length of hospital stay. Features that demonstrated significant distinctions in univariate analysis were incorporated into the development of ML models, which were subsequently subjected to cross-validation. The research then undertook a comparative assessment of the ML models' performance and identified the risk factors associated with eLOS.

Research results

Incorporating a cohort of 763 patients, of which 380 experienced eLOS, the study evaluated the predictive performance of various ML models, with decision tree random forest, and eXtreme Gradient Boosting models emerging as the most robust. Additionally, the artificial neural network model demonstrated commendable results. Following cross-validation, the support vector machine and logistic regression models displayed superior predictive capabilities. Key predictors for eLOS encompassed delayed surgery, D-dimer levels, ASA classification, type of surgery, and gender.

Research conclusions

The application of ML yielded exceptional accuracy in forecasting eLOS among geriatric hip fracture patients. Notably, the study identified significant risk factors, including delayed surgery, D-dimer levels, ASA classification, surgical procedure type, and gender. This valuable insight has the potential to assist clinicians in optimizing resource allocation to meet patient demands more effectively.

Research perspectives

Future research in ML applications for predicting eLOS in geriatric hip fracture patients will likely focus on refining models, integrating them into clinical practice, ensuring interpretability, and addressing ethical and practical considerations to enhance the utility and impact of these predictive tools in healthcare.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to Dr. Rui for the proofreading of this manuscript.

FOOTNOTES

Author contributions: Tian CW and Chen XX contributed equally to this work; Tian CW and Chen XX designed this study and wrote the manuscript; Zhu HY performed the experiments and analyzed the data; Tian CW and Zhu HY collected the clinical data; Shi L, Dai GC and Chen H provided technical support; Rui YF provided the idea and revised and proofread the paper; All the authors have read and approved the final manuscript.

Supported by Winfast Charity Foundation for Financial Support, No. YL20220525.

Institutional review board statement: This study was reviewed and approved by the Ethics Committee of the Zhongda Hospital, Affiliated to Southeast University (2022ZDSYLL183-P01).

Informed consent statement: All study participants or their legal guardian provided informed written consent about personal and medical data collection prior to study enrolment.

Conflict-of-interest statement: The authors declare no competing financial interests or conflicts of interest that could have influenced the design, data collection, analysis, interpretation, or publication of this study.

Data sharing statement: Requests for data access should be directed to the corresponding author at ruiyunfeng@126.com.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

Country/Territory of origin: China

ORCID number: Yun-Feng Rui 0000-0001-9019-5531.

S-Editor: Qu XL

L-Editor: A

P-Editor: Zhao S

REFERENCES

- 1 Veronese N, Maggi S. Epidemiology and social costs of hip fracture. *Injury* 2018; **49**: 1458-1460 [PMID: 29699731 DOI: 10.1016/j.injury.2018.04.015]
- 2 Aasvang EK, Luna IE, Kehlet H. Challenges in postdischarge function and recovery: the case of fast-track hip and knee arthroplasty. *Br J Anaesth* 2015; **115**: 861-866 [PMID: 26209853 DOI: 10.1093/bja/aeu257]
- 3 Soffin EM, YaDeau JT. Enhanced recovery after surgery for primary hip and knee arthroplasty: a review of the evidence. *Br J Anaesth* 2016; **117**: iii62-iii72 [PMID: 27940457 DOI: 10.1093/bja/aew362]
- 4 Rosman M, Rachminov O, Segal O, Segal G. Prolonged patients' In-Hospital Waiting Period after discharge eligibility is associated with increased risk of infection, morbidity and mortality: a retrospective cohort analysis. *BMC Health Serv Res* 2015; **15**: 246 [PMID: 26108373 DOI: 10.1186/s12913-015-0929-6]
- 5 Paton F, Chambers D, Wilson P, Eastwood A, Craig D, Fox D, Jayne D, McGinnes E. Effectiveness and implementation of enhanced recovery after surgery programmes: a rapid evidence synthesis. *BMJ Open* 2014; **4**: e005015 [PMID: 25052168 DOI: 10.1136/bmjopen-2014-005015]
- 6 Slim K. Fast-track surgery: the next revolution in surgical care following laparoscopy. *Colorectal Dis* 2011; **13**: 478-480 [PMID: 21435146 DOI: 10.1111/j.1463-1318.2011.02589.x]
- 7 Xie T, Ma B, Li Y, Zou J, Qiu X, Chen H, Wang C, Rui Y. [Research status of the enhanced recovery after surgery in the geriatric hip fractures]. *Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi* 2018; **32**: 1038-1046 [PMID: 30238732 DOI: 10.7507/1002-1892.201712083]
- 8 Schneider AM, Denyer S, Brown NM. Risk Factors Associated With Extended Length of Hospital Stay After Geriatric Hip Fracture. *J Am Acad Orthop Surg Glob Res Rev* 2021; **5**: e21.00073 [PMID: 33945514 DOI: 10.5435/JAAOSGlobal-D-21-00073]
- 9 Deo RC. Machine Learning in Medicine. *Circulation* 2015; **132**: 1920-1930 [DOI: 10.1161/CIRCULATIONAHA.115.001593]
- 10 De Silva T, Hess K, Grisso P, Thavikulwat AT, Wiley H, Keenan TDL, Chew EY, Jeffrey BG, Cukras CA. Deep Learning-Based Modeling of the Dark Adaptation Curve for Robust Parameter Estimation. *Transl Vis Sci Technol* 2022; **11**: 40 [PMID: 36315120 DOI: 10.1167/tvst.11.10.40]
- 11 Karan Aggarwal, Maad M. Mijwil, Sonia, Abdel-Hameed Al-Mistarehi, Safwan Alomari, Murat Gök, Anas M. Zein Alaabdin, Safaa H. Abdulrhman. Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning. *Iraqi Journal for Computer Science and Mathematics* 2022; **115**-123 [DOI: 10.52866/ijcsm.2022.01.01.013]
- 12 Mijwil MM, Aggarwal K. A diagnostic testing for people with appendicitis using machine learning techniques. *Multimed Tools Appl* 2022; **81**: 7011-7023 [PMID: 35095329 DOI: 10.1007/s11042-022-11939-8]
- 13 Oosterhoff JHF, Savelberg ABMC, Karhade AV, Gravesteijn BY, Doornberg JN, Schwab JH, Heng M. Development and internal validation of a clinical prediction model using machine learning algorithms for 90 day and 2 year mortality in femoral neck fracture patients aged 65 years or above. *Eur J Trauma Emerg Surg* 2022; **48**: 4669-4682 [PMID: 35643788 DOI: 10.1007/s00068-022-01981-4]
- 14 Shtar G, Rokach L, Shapira B, Nissan R, Hershkovitz A. Using Machine Learning to Predict Rehabilitation Outcomes in Postacute Hip Fracture Patients. *Arch Phys Med Rehabil* 2021; **102**: 386-394 [PMID: 32949551 DOI: 10.1016/j.apmr.2020.08.011]
- 15 Rui Y, Qiu X, Zou J, Xie T, Ma B, Lu P, Li Y, Liu S, Jin J, Deng C, Cui Y, Wang X, Ma M, Ren L, Yang Y, Wang C, Chen H. [Clinical application of multidisciplinary team co-management in geriatric hip fractures]. *Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi* 2019; **33**: 1276-1282 [PMID: 31544439 DOI: 10.7507/1002-1892.201905017]
- 16 Poitras S, Au K, Wood K, Dervin G, Beaulé PE. Predicting hospital length of stay and short-term function after hip or knee arthroplasty: are both performance and comorbidity measures useful? *Int Orthop* 2018; **42**: 2295-2300 [PMID: 29453585 DOI: 10.1007/s00264-018-3833-y]
- 17 Konda SR, Johnson JR, Kelly EA, Chan J, Lyon T, Egol KA. Can We Accurately Predict Which Geriatric and Middle-Aged Hip Fracture Patients Will Experience a Delay to Surgery? *Geriatr Orthop Surg Rehabil* 2020; **11**: 2151459320946021 [PMID: 32821470 DOI: 10.1177/2151459320946021]
- 18 Lott A, Haglin J, Belayneh R, Konda SR, Egol KA. Admitting Service Affects Cost and Length of Stay of Hip Fracture Patients. *Geriatr Orthop Surg Rehabil* 2018; **9**: 2151459318808845 [PMID: 30479850 DOI: 10.1177/2151459318808845]
- 19 Mathew PJ, Jehan F, Kulvatunyou N, Khan M, O'Keeffe T, Tang A, Gries L, Hamidi M, Zakaria ER, Joseph B. The burden of excess length of stay in trauma patients. *Am J Surg* 2018; **216**: 881-885 [PMID: 30082028 DOI: 10.1016/j.amjsurg.2018.07.044]
- 20 Diaz-Ledezma C, Mardones R. Predicting Prolonged Hospital Stays in Elderly Patients With Hip Fractures Managed During the COVID-19 Pandemic in Chile: An Artificial Neural Networks Study. *HSS J* 2023; **19**: 205-209 [PMID: 37051613 DOI: 10.1177/15563316221120582]
- 21 Gabriel RA, Sharma BS, Doan CN, Jiang X, Schmidt UH, Vaida F. A Predictive Model for Determining Patients Not Requiring Prolonged Hospital Length of Stay After Elective Primary Total Hip Arthroplasty. *Anesth Analg* 2019; **129**: 43-50 [PMID: 30234533 DOI: 10.1213/ANE.00000000000003798]
- 22 Pincus D, Wasserstein D, Ravi B, Huang A, Paterson JM, Jenkinson RJ, Kreder HJ, Nathens AB, Wodchis WP. Medical Costs of Delayed Hip Fracture Surgery. *J Bone Joint Surg Am* 2018; **100**: 1387-1396 [PMID: 30106820 DOI: 10.2106/JBJS.17.01147]
- 23 Hecht G, Slee CA, Goodell PB, Taylor SL, Wolinsky PR. Predictive Modeling for Geriatric Hip Fracture Patients: Early Surgery and Delirium Have the Largest Influence on Length of Stay. *J Am Acad Orthop Surg* 2019; **27**: e293-e300 [PMID: 30358636 DOI: 10.5435/JAAOS-D-17-00447]
- 24 Kristan A, Omahen S, Tosounidis TH, Cimerman M. When does hip fracture surgery delay affects the length of hospital stay? *Eur J Trauma Emerg Surg* 2022; **48**: 701-708 [PMID: 33388817 DOI: 10.1007/s00068-020-01565-0]
- 25 Niikura T, Sakai Y, Lee SY, Iwakura T, Nishida K, Kuroda R, Kurosaka M. D-dimer levels to screen for venous thromboembolism in patients with fractures caused by high-energy injuries. *J Orthop Sci* 2015; **20**: 682-688 [PMID: 25797331 DOI: 10.1007/s00776-015-0711-y]
- 26 Garcia AE, Bonnaig JV, Yoneda ZT, Richards JE, Ehrenfeld JM, Obremskey WT, Jahangir AA, Sethi MK. Patient variables which may predict length of stay and hospital costs in elderly patients with hip fracture. *J Orthop Trauma* 2012; **26**: 620-623 [PMID: 22832431 DOI: 10.1097/BOT.0b013e3182695416]

- 27 **Haentjens P**, Magaziner J, Colón-Emeric CS, Vanderschueren D, Milisen K, Velkeniers B, Boonen S. Meta-analysis: excess mortality after hip fracture among older women and men. *Ann Intern Med* 2010; **152**: 380-390 [PMID: [20231569](#) DOI: [10.7326/0003-4819-152-6-201003160-00008](#)]
- 28 **Hawkes WG**, Wehren L, Orwig D, Hebel JR, Magaziner J. Gender differences in functioning after hip fracture. *J Gerontol A Biol Sci Med Sci* 2006; **61**: 495-499 [PMID: [16720747](#) DOI: [10.1093/gerona/61.5.495](#)]
- 29 **Castelli A**, Daidone S, Jacobs R, Kasteridis P, Street AD. The Determinants of Costs and Length of Stay for Hip Fracture Patients. *PLoS One* 2015; **10**: e0133545 [PMID: [26204450](#) DOI: [10.1371/journal.pone.0133545](#)]
- 30 **Lefavre KA**, Macadam SA, Davidson DJ, Gandhi R, Chan H, Broekhuysen HM. Length of stay, mortality, morbidity and delay to surgery in hip fractures. *J Bone Joint Surg Br* 2009; **91**: 922-927 [PMID: [19567858](#) DOI: [10.1302/0301-620X.91B7.22446](#)]
- 31 **Obermeyer Z**, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med* 2016; **375**: 1216-1219 [PMID: [27682033](#) DOI: [10.1056/NEJMp1606181](#)]
- 32 **Cha Y**, Kim JT, Park CH, Kim JW, Lee SY, Yoo JI. Artificial intelligence and machine learning on diagnosis and classification of hip fracture: systematic review. *J Orthop Surg Res* 2022; **17**: 520 [PMID: [36456982](#) DOI: [10.1186/s13018-022-03408-7](#)]
- 33 **Kang YJ**, Yoo JI, Cha YH, Park CH, Kim JT. Machine learning-based identification of hip arthroplasty designs. *J Orthop Translat* 2020; **21**: 13-17 [PMID: [32071870](#) DOI: [10.1016/j.jot.2019.11.004](#)]
- 34 **Lalehzarian SP**, Gowd AK, Liu JN. Machine learning in orthopaedic surgery. *World J Orthop* 2021; **12**: 685-699 [PMID: [34631452](#) DOI: [10.5312/wjo.v12.i9.685](#)]
- 35 **Forssten MP**, Bass GA, Ismail AM, Mohseni S, Cao Y. Predicting 1-Year Mortality after Hip Fracture Surgery: An Evaluation of Multiple Machine Learning Approaches. *J Pers Med* 2021; **11** [PMID: [34442370](#) DOI: [10.3390/jpm11080727](#)]
- 36 **Galassi A**, Martín-Guerrero JD, Villamor E, Monserrat C, Rupérez MJ. Risk Assessment of Hip Fracture Based on Machine Learning. *Appl Bionics Biomech* 2020; **2020**: 8880786 [PMID: [33425008](#) DOI: [10.1155/2020/8880786](#)]
- 37 **Luo X**, Wen X, Zhou M, Abusorrah A, Huang L. Decision-Tree-Initialized Dendritic Neuron Model for Fast and Accurate Data Classification. *IEEE Trans Neural Netw Learn Syst* 2022; **33**: 4173-4183 [PMID: [33729951](#) DOI: [10.1109/TNNLS.2021.3055991](#)]
- 38 **Wang R**, Zhang J, Shan B, He M, Xu J. XGBoost Machine Learning Algorithm for Prediction of Outcome in Aneurysmal Subarachnoid Hemorrhage. *Neuropsychiatr Dis Treat* 2022; **18**: 659-667 [PMID: [35378822](#) DOI: [10.2147/NDT.S349956](#)]
- 39 **Chen T**, G Carlos. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2016 [DOI: [10.1145/2939672.2939785](#)]
- 40 **Hou N**, Li M, He L, Xie B, Wang L, Zhang R, Yu Y, Sun X, Pan Z, Wang K. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *J Transl Med* 2020; **18**: 462 [PMID: [33287854](#) DOI: [10.1186/s12967-020-02620-5](#)]
- 41 **Noh B**, Youm C, Goh E, Lee M, Park H, Jeon H, Kim OY. XGBoost based machine learning approach to predict the risk of fall in older adults using gait outcomes. *Sci Rep* 2021; **11**: 12183 [PMID: [34108595](#) DOI: [10.1038/s41598-021-91797-w](#)]
- 42 **Chen CY**, Chen YF, Chen HY, Hung CT, Shi HY. Artificial Neural Network and Cox Regression Models for Predicting Mortality after Hip Fracture Surgery: A Population-Based Comparison. *Medicina (Kaunas)* 2020; **56** [PMID: [32438724](#) DOI: [10.3390/medicina56050243](#)]
- 43 **Tzzy-Chyang Lu**, Gwo-Ruey Yu, Jyh-Ching Juang. Quantum-based algorithm for optimizing artificial neural networks. *IEEE Trans Neural Netw Learn Syst* 2013; **24**: 1266-1278 [PMID: [24808566](#) DOI: [10.1109/TNNLS.2013.2249089](#)]
- 44 **Lapuerta P**, Azen SP, LaBree L. Use of neural networks in predicting the risk of coronary artery disease. *Comput Biomed Res* 1995; **28**: 38-52 [PMID: [7614823](#) DOI: [10.1006/cbmr.1995.1004](#)]
- 45 **Klemm C**, Yeo I, Cohen-Levy WB, Melnic CM, Habibi Y, Kwon YM. Artificial Neural Networks Can Predict Early Failure of Cementless Total Hip Arthroplasty in Patients With Osteoporosis. *J Am Acad Orthop Surg* 2022; **30**: 467-475 [PMID: [35202042](#) DOI: [10.5435/JAAOS-D-21-00775](#)]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA

Telephone: +1-925-3991568

E-mail: bpgoffice@wjgnet.com

Help Desk: <https://www.f6publishing.com/helpdesk>

<https://www.wjgnet.com>

