87059_Auto_Edited.docx

Case Control Study

Machine learning applications for the prediction of eLOS in geriatric hip fracture patients: a case control study

Machine learning for hip fracture

Chuwei Tian, Xiangxu Chen, Liu Shi, Huanyi Zhu, Guang-Chun Dai, Hui Chen, Yun-Feng Rui

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

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.

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 LOS. 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 (DT), random forest (RF), and eXtreme Gradient Boosting (XGB) models demonstrated the most robust performance. Notably, the artificial neural network (ANN) model also exhibited impressive results. After cross-validation, the support vector machine (SVM) and logistic regression (LR) 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

Machine learning 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; ERAS; risk factors

Tian C, Chen X, Shi L, Zhu H, Dai GC, Chen H, Rui YF. Machine learning applications for the prediction of eLOS in geriatric hip fracture patients: a case control study. *World J Orthop* 2023; In press

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 ERAS 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.

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* developed a logistic regression model to identify risk factors for extended length of stay (eLOS), offering new insights for optimizing treatment for hip fracture patients [8]. 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, first proposed by Arthur Samuel in 1959 [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 *et al* enhanced the ML-based estimation method for detecting acute appendicitis in individuals, achieving high accuracy [12]. In the context of hip fractures among geriatric individuals, Shtar *et al* and Oosterhoff *et al* established ML models to predict prognosis and mortality, enhancing clinician decision-making ability [13, 14]. With the establishment

of a clinical database for elderly hip fracture patients and advancements in machine learning algorithms, predicting the eLOS for this patient group through machine learning 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,

3. Hospitalization at our centre.

Exclusion Criteria:

1. Admission with multiple fractures, pathological fracture, or fractures around the prosthesis,

2. Receipt of conservative treatment due to severe comorbidities,

3. Presence of missing data.

The enrolled patients had a median hospital stay of 9.5 days. Based on this median length of stay, the patients were retrospectively divided into two groups: noneLOS (length of stay ≤ 9.5 days, $n = 383$, 50.2%) and eLOS (length of stay > 9.5 days, $n = 380$, 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 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 (HHA), 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 (BMI), 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 ≤ 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.

Machine learning and statistical analysis

In the study, normally distributed data are presented as the means and standard deviations (SDs). Nonnormally distributed variables were expressed as medians along with their interquartile ranges (IQRs). 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 machine learning (ML) model.

The predictive eLOS ML model was established according to the selected features, including basic algorithms for logistic regression (LR), decision tree (DT), random forest (RF), support vector machine (SVM), naïve Bayes (NB), K-nearest neighbour (KNN), 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 confusion matrix, 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 noneLOS (n = 383) and eLOS (n = 380) groups based on length of stay. 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, AST and D-dimer levels and aortic velocity (AV) 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.

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) model showed good performance after cross-validation.

DISCUSSION

This study aimed to develop machine learning (ML) models for predicting extended length of stay (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* showed that patients who underwent surgery 1 day after admission had a longer postoperative stay [22]. Hecht *et al* developed a predictive model for LOS that showed that the prevention of reduced time to surgery was a significant predictor of reduced LOS. In addition, the authors suggested that ASA classification was a stronger predictor of LOS than the Charlson Comorbidity Index (CCI) [23]. Similar findings were shown by Kristan *et al*, who found that the eLOS was associated with delayed surgery, ASA classification, anticoagulant therapy and surgery type [24]. 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* 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 [26]. A meta-analysis by Haentjens *et al* 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 [27, 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 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*, who used ML to predict 1-year mortality after hip fracture surgery, and by Galassi *et al*, who used ML to assess hip fracture risk [35, 36]. 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 decision trees, 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* used XGBoost to predict 30-day mortality for the Medical Information Mart for Intensive Care III patients with sepsis-3 and obtained high accuracy [40]. Noh *et al* similarly achieved good accuracy in identifying the optimal features of gait parameters to predict fall risk among older adults by XGBoost [41]. 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* showed that the ANN was more accurate than Cox regression in predicting mortality after hip fracture surgery [42]. 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 Klemt *et al* demonstrated its good application in binary data [45]. 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 machine learning (ML) models for predicting extended length of stay (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 length of stay and allocate medical resources effectively, aligning with the enhanced recovery after surgery (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 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 machine learning 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 machine learning 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 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 machine learning models to forecast extended length of stay (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

+ADw-html+AD4APA-p+AD4-This research endeavors to construct machine learning models for the purpose of forecasting extended length of stay (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 machine learning model for the prediction of extended length of stay (eLOS) in patients with hip fractures. Significantly, factors such as delayed surgical intervention, elevated D-dimer levels, ASA classification, surgical procedure type, and gender exhibited notable associations with eLOS. The integration of machine learning 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.+ADw-/p+AD4APA-/html+AD4-

Research methods

+ADw-html+AD4APA-p+AD4-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 (LOS). Features that demonstrated significant distinctions in univariate analysis were incorporated into the development of machine learning models, which were subsequently subjected to cross-validation. The research then undertook a

comparative assessment of the machine learning models' performance and identified the risk factors associated with extended length of stay (eLOS).+ADw-/p+AD4APA-/html+AD4-

Research results

+ADw-html+AD4APA-p+AD4-Incorporating a cohort of 763 patients, of which 380 experienced extended length of stay (eLOS), the study evaluated the predictive performance of various machine learning models, with decision tree (DT), random forest (RF), and eXtreme Gradient Boosting (XGB) models emerging as the most robust. Additionally, the artificial neural network (ANN) model demonstrated commendable results. Following cross-validation, the support vector machine (SVM) and logistic regression (LR) models displayed superior predictive capabilities. Key predictors for eLOS encompassed delayed surgery, D-dimer levels, American Society of Anaesthesiologists (ASA) classification, type of surgery, and gender.+ADw-/p+AD4APA-/html+AD4-

Research conclusions

The application of machine learning yielded exceptional accuracy in forecasting extended length of stay (eLOS) among geriatric hip fracture patients. Notably, the study identified significant risk factors, including delayed surgery, D-dimer levels, American Society of Anesthesiologists (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 machine learning 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

87059_Auto_Edited.docx

ORIGINALITY REPORT

9%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet	127 words — **3%**
2	Chuwei Tian, Huanyi Zhu, Liu Shi, Xiangxu Chen, Tian Xie, Yunfeng Rui. "Is There a "Black Friday" for Geriatric Hip Fracture Surgery?", Orthopaedic Surgery, 2023 Crossref	75 words — **2%**
3	www.frontiersin.org Internet	25 words — **1%**
4	www.biorxiv.org Internet	22 words — **1%**
5	downloads.hindawi.com Internet	20 words — **1%**
6	www.mdpi.com Internet	19 words — **< 1%**
7	Carmelo Agostino Di Franco, Sandro Bellavia, Carmela Emmanuele, Genziana Buttafuoco et al. "A Rare Case of Testicular Plasmacytoma as First Presentation in a Monorchid Patient Without Previous History of Multiple Myeloma: Diagnosis, Management and Review of Literature", World Journal of Nephrology and Urology, 2023 Crossref	14 words — **< 1%**

8 academic-accelerator.com
Internet

14 words — < 1%

9 www.nature.com
Internet

13 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 12 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 12 WORDS