International Journal of Medical Informatics 155 (2021) 104572

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

The application of machine learning algorithms in predicting the length of stay following femoral neck fracture

Hao Zhong [a,1], Bingpu Wang [b,1], Dawei Wang [a], Zirui Liu [c], Cong Xing [a], Yu Wu [d], Qiang Gao [a], Shibo Zhu [a], Haodong Qu [a], Zeyu Jia [a], Zhigang Qu [e,*], Guangzhi Ning [a,*], Shiqing Feng [a,*]

[a] *International Science and Technology Cooperation Base of Spinal Cord Injury, Tianjin Key Laboratory of Spine and Spinal Cord Injury, Department of Orthopedics, Tianjin Medical University General Hospital, 154 Anshan Road, Heping District, Tianjin, China*

[b] *State Key Laboratory of Precision Measurement Technology and Instrument, School of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China*

[c] *Texas A&M University, United States*

[d] *Department of Orthopedics, The First People's Hospital of Yichang, YiChang, Hubei Province, China*

[e] *College of electronic information and automation, Tianjin University of Science and Technology, Tianjin, China*

ARTICLE INFO

Keywords:
Machine learning
Length of stay
Femoral neck fracture
Intravenous fluid management

ABSTRACT

Purpose: Femoral neck fracture is a frequent cause of hospitalization, and length of stay is an important marker of hospital cost and quality of care provided. As an extension of traditional statistical methods, machine learning provides the possibility of accurately predicting the length of hospital stay. The aim of this paper is to retrospectively identify predictive factors of the length of hospital stay (LOS) and predict the postoperative LOS by using machine learning algorithms.

Method: Based on the admission and perioperative data of the patients, linear regression was used to analyze the predictive factors of the LOS. Multiple machine learning models were developed, and the performance of different models was compared.

Result: Stepwise linear regression showed that preoperative calcium level ($P = 0.017$) and preoperative lymphocyte percentage ($P = 0.007$), in addition to intraoperative bleeding ($p = 0.041$), glucose and sodium chloride infusion after surgery ($P = 0.019$), Charlson Comorbidity Index ($p = 0.007$) and BMI ($P = 0.031$), were significant predictors of LOS. The best performing model was the principal component regression (PCR) with an optimal MAE (1.525) and a proportion of prediction error within 3 days of 90.91%.

Conclusion: Excessive intravenous glucose and sodium chloride infusion after surgery, preoperative hypocalcemia, preoperative high percentages of lymphocytes, excessive intraoperative bleeding, lower BMI and higher CCI scores were related to prolonged LOS by using linear regression. Machine learning could accurately predict the postoperative LOS. This information allows hospital administrators to plan reasonable resource allocation to fulfill demand, leading to direct care quality improvement and more reasonable use of scarce resources.

1. Introduction

Femoral neck fractures occur most commonly in aged individuals with osteoporosis and impose a substantial financial burden on the public healthcare system and society. Femoral neck fractures, which account for approximately half of hip fractures, are mainly caused by trauma, such as falling, and the mental and physical health of hip fracture patients are negatively impacted by the injury [1]. In China, the direct medical cost of hip fractures will increase to 4.063 billion dollars in 2050, more than doubling the cost in 2018 [2,3].

Femoral neck fracture is a frequent cause of hospitalization in the orthopedic department. Additionally, LOS has been used as a common surrogate marker for hospital cost and quality of care provided [4]. There are many costs associated with femoral neck fracture, and hospitalization is an important contributor to this cost. LOS could reflect the number of medical resources used. The consequence of prolonged LOS is an increased risk of complications and excessive use of resources [5]. By identifying excessively short or prolonged LOS, earlier and hence more

* Corresponding authors.
E-mail addresses: zhigangqu@tust.edu.cn (Z. Qu), ningguangzhi@foxmail.com (G. Ning), sqfeng@tmu.edu.cn (S. Feng).

[1] The first two authors (Hao Zhong and Bingpu Wang) contributed equally to this work.

https://doi.org/10.1016/j.ijmedinf.2021.104572
Received 3 March 2021; Received in revised form 7 September 2021; Accepted 8 September 2021
Available online 13 September 2021

effective intervention strategies can be identified. Successful prediction of the LOS will bring substantial benefits to both hospitals and patients [6].

A variety of studies have tried to predict the LOS or determine the factors related to the LOS. A. E. Garcia et al. collected 660 complete records and found that ASA classification was a reliable predictor of LOS in hip fracture patients [7]. John E et al. examined several factors and found that increased age and chronic cognitive impairment/dementia could prolong the LOS [8]. Because of the complexity of the patient's pathophysiological mechanism and hospital management factors, constructing an accurate algorithm is difficult. Traditional statistical methods are standard when developing prediction models but are not as accurate as machine learning [9]. Driven by the unparalleled wealth of data and the rapid development of processing power, more attention has been given to machine learning. As an extension of traditional statistical methods, machine learning is widely used in daily life. The applications of machine learning in medicine have been pioneering [10].

The aim of this study was to analyze predictors of the LOS of patients with femoral neck fracture. Multiple machine learning models were developed to predict the postoperative LOS(PLOS) and the performance of different models was compared.

2. Materials and methods

2.1. Data collection

The research was conducted at Tianjin Medical University General Hospital, the largest regional teaching hospital in Tianjin. We included all patients who were diagnosed with femoral neck fracture and underwent surgery at the orthopedic department between July 2017 and Dec 2019. Patients were identified by searching the electronic medical records. These data elements were manually entered by professional orthopedists, and each record was inspected by medical teams. Individuals with 20% missing data, with an LOS greater than 60 days, and who died before discharge were excluded from the study. A total of 182 inpatient records fulfilled the inclusion criteria. Data were collected using a structured form. Patients routinely collected preoperative laboratory values, including electrolytes, creatinine, liver function tests, blood counts and intraoperative situations, were collected according to the related literature (as shown in Table 1). Variables are summarized as the mean ± SD for numerical variables and proportion for categorical variables. Intravenous fluid administration after surgery on the day of surgery was divided into four groups: 1: 0.9% sodium chloride injection (the unit was 100 ml); 2: compound amino acid injection (the unit was 500 ml); 3: glucose and sodium chloride injection (the unit was 500 ml); and 4: alanyl glutamine (the unit was 20 g). Thirty-eight variables that were divided into numerical or categorical variables were extracted. Dummy variables were used to normalize the categorical variables that do not have a natural rank ordering.

2.2. Statistical methods and machine learning algorithms

2.2.1. Statistical methods

Data were entered, coded, and analyzed using IBM SPSS version 23.0 for Windows. Subsequently, stepwise multiple linear regression analysis was used to analyze factors associated with the LOS when independent variables were categorical. Statistical significance is noted by asterisks: *, P value of <0.05; **, P value of <0.01.

2.2.2. Machine learning algorithms

In recent years, with the rapid development of artificial intelligence, there have been many available algorithms. The choice of prediction algorithm will directly affect the final prediction result, which is the key to the performance of the model [11]. In this study, three algorithms were selected to establish the prediction model.

Table 1
Demographic information and clinical characteristics.

Variables	
Length of stay, mean ± SD	16.6 ± 6.9
Postoperative length of stay, mean ± SD	11.8 ± 6.1
Age, years, mean ± SD	74.1 ± 10.6
Sex, %(n)	
Male	30.8 (56)
Female	69.2 (126)
Body mass index (BMI), mean	23.3 ± 4.4
Drinking History, %(n)	14.7 (26)
Smoking History, %(n)	10.2 (18)
Hypertension, %(n)	51.1 (94)
Charlson Comorbidity Index	
0	45.6 (83)
1	29.1 (53)
2	8.8 (16)
3	10.4 (19)
4	4.4 (8)
5	1.1(2)
6	0.5(1)
Time from fracture to hospitalization, hours, mean ± SD	77.2 ± 159.0
Surgery Category, %(n)	
Open Reduction Internal Fixation	7.7 (14)
Hemiarthroplasty	45.1 (82)
Total Hip Arthroplasty	46.7 (85)
Other	0.5 (1)
American Society of Anesthesiologists (ASA) classification	
1	0.5 (1)
2	34.1 (62)
3	63.2 (115)
4	2.2 (4)
Anesthesia Time, minutes, mean ± SD	145.6 ± 31.6
Intraoperative Bleeding, ml, mean ± SD	108.3 ± 75.2
Blood Transfusion, %(n)	2.7 (5)
Intraoperative Colloidal Fluid, ml, mean	266.3
Intraoperative Crystal Fluid, ml, mean	1054.5
Intraoperative Urine Output, ml, mean	267.1
Anesthesia Category, % (n)	
Spinal Anesthesia	70.9 (129)
General Anesthesia	27.5 (50)
Other	1.6 (3)
Intraoperative urine output, ml, mean	263.8
Preoperative Hypertension	66.5 (121)
Postoperative Hypertension	38.5% (70)
Intravenous fluid administration after surgery on the day of surgery	
0.9% sodium chloride injection, ml, mean	558.1
compound amino acid injection, ml, mean	65.9
glucose and sodium chloride injection, ml, mean	693.0
alanyl glutamine, ml, mean	21.4
Acid inhibitory drug, % (n)	84.6 (154)

2.2.2.1. Back propagation (BP) neural network. The BP neural network model is composed of three layers, including the input layer, hidden layer, and output layer. The input layer is responsible for reading the patient data in the sample. Each node represents the patient's medical record information. There are 38 nodes in total. The main function of the hidden layer is to extract features from the input layer of the previous layer, and the output of neurons in the hidden layer is as follows:

$$y_i = f_i(\omega_{(x_1)i}x_1 + \omega_{(x_2)i}x_2 + \dots + \omega_{(x_j)i}x_j) \tag{1}$$

where y_i is the output of the i-th node of the hidden layer, $\omega_{(x_j)i}$ is the weight of the j-th input node to the i-th node of the hidden layer, and$x_j$is the j-th input node.

Then, the relationship between the hidden layer and the output layer is established to complete the positive propagation of the neural network. Next, (2) the back propagation of the neural network will be carried out, and the error information will be fed back from the output layer to the hidden layer. Then, the error will be modified by the gradient descent method to make the neural network stable.

$$w'_{(x_j)i} = w_{(x_j)i} + \eta\Delta\frac{df_i(e)}{de}x_j \quad (2)$$

where η is the learning rate and$\Delta$is the error of back propagation of the latter layer.

In this way, the renewal of the connection weight is completed, and then forward propagation, calculation error, and back propagation are performed to repair the weight layer by layer, and the cycle repeats.

2.2.2.2. Support vector Machine (SVM). When SVM is applied to regression problems, the basic idea is to find an optimal classification surface to minimize the error of the optimal classification surface for all training sample distances. The hyperplane is divided in the sample space by the following linear equation:

$$w^Tx + b = 0, w^T = (w_1, w_2, \dots w_3) \quad (3)$$

where x is the input parameter of the training set.

Therefore, the distance from any point in the sample to the hyperplane is as follows:

$$r = \frac{|w^Tx + b|}{\|w\|}$$
$$s.t\begin{cases} w^Tx_i + b \geqslant +1 \\ w^Tx_i + b \leqslant -1 \end{cases} \quad (4)$$

Several training samples points closest to the hyperplane make the inequality true. The sum of the distances of two heterogeneous support vectors to the hyperplane is the interval. Identifying*w*and*b*make the inequality true. At this point, r is at its maximum:

$$max(r) = max(\frac{2}{\|w\|}) \quad (5)$$

The model can be obtained by solving *w* and *b*.

2.2.2.3. Principle component regression (PCR). The principle of PCR is to replace the original variable with a few new variables, which can reflect the data information of the original variable as much as possible. In these circumstances, the resulting new variables can expel the superimposed information in the original variables, and the new variables are orthogonal to each other. First a variable that satisfies the following is constructed:

$$P_1 = X\|t\|, \|t\| = 1 \quad (6)$$

where X is the matrix containing all the input sample information, and t is a unit vector.

From the perspective of probability statistics, the greater the variance of the variable, the more information is contained in the variable, so it is converted to finding the maximum variance of the variable P_1. The variance of P_1 is calculated as follows:

$$Var(P_1) = \frac{1}{n}\|P_1\|^2 = \frac{1}{n}tX'Xt = t'Vt \;\; \mathrm{V} \;=\; \frac{1}{n}X'X \quad (7)$$

Now the Lagrangian is constructed.

$$L = t'Vt - \lambda(t't - 1) \quad (8)$$

where λ is the Lagrangian coefficient. Partial derivatives of L with respect to λ and t are calculated and set to zero, then

$$\begin{cases} \frac{\partial L}{\partial t} = 2Vt - 2\lambda t = 0 \\ \frac{\partial L}{\partial \lambda} = -(t't - 1) = 0 \end{cases} \quad \mathrm{Vt} \;=\; \lambda\mathrm{t} \quad (9)$$

From (9), one can perceive that t is a standardized eigenvector of V, and λ is its homologous eigenvalue.

$$Var(P_1) = t'Vt \;=\; t'\lambda t = \lambda t't = \lambda \quad (10)$$

In other words, the required t is the normalized eigenvector corresponding to the maximum value of the eigenvalue λ of the matrix V, and the homologous construction variable P_1 is named the first principal component. By analogy, multiple principal components of X can be determined.

The total amount of information included in the first m principal components is as follows:

$$\sum_{i=1}^{m} Var(P_i) = \sum_{i=1}^{m}\lambda_i \quad (11)$$

In this paper, principal component analysis was performed on the input data from the training samples, and the result is shown in Fig. 1. The 38 variables in the data set were converted into 6 comprehensive indicators. PCR is a multiple linear regression analysis based on principal component analysis. Fig. 2 illustrating the sampling procedure and steps of the proposed method.

2.2.2.4. Data preprocessing. In order to give full play to the prediction function of each prediction model and improve the prediction accuracy, it is very necessary to preprocess the data. In the existing samples, part of the input features contains missing values, for which we respectively adopted two often used solutions: case deletion and mean filling. In order to accelerate the training speed of the model and prevent some neurons from reaching saturation state in the training process, the data was normalized. In this study, the input data were finally normalized to [0,1] through experimental verification. The specific method is to first calculate the maximum value Xmax and minimum value Xmin in the sample data, and X' and X respectively represent the data before and after normalization. Then the normalization publicity of the sample data is

$$X = \frac{X' - X_{min}}{X_{max} - X_{min}} \quad (12)$$

2.2.2.5. Parameter Determination in the Prediction Model. **BP neural network mainly contains the following parameters:**

- Network Layer Design

Without limiting the number of hidden neurons, the BP network with only one hidden layer in three layers can realize the nonlinear mapping relationship. In the case of relatively few samples, one hidden network can be selected, so one hidden layer is selected in this study.

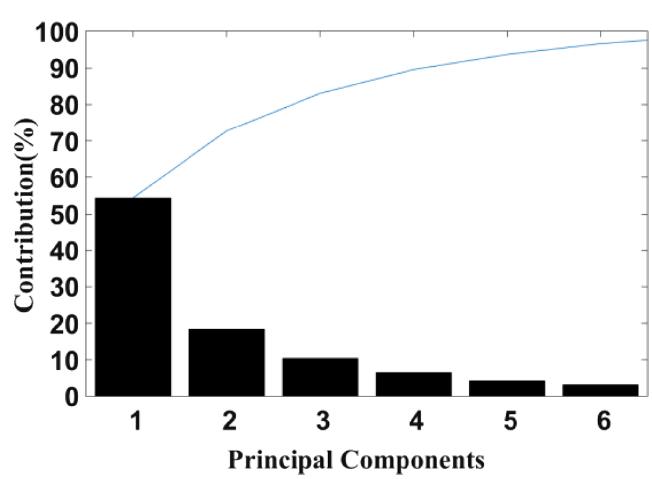

Fig. 1. PCR principal component contribution analysis.

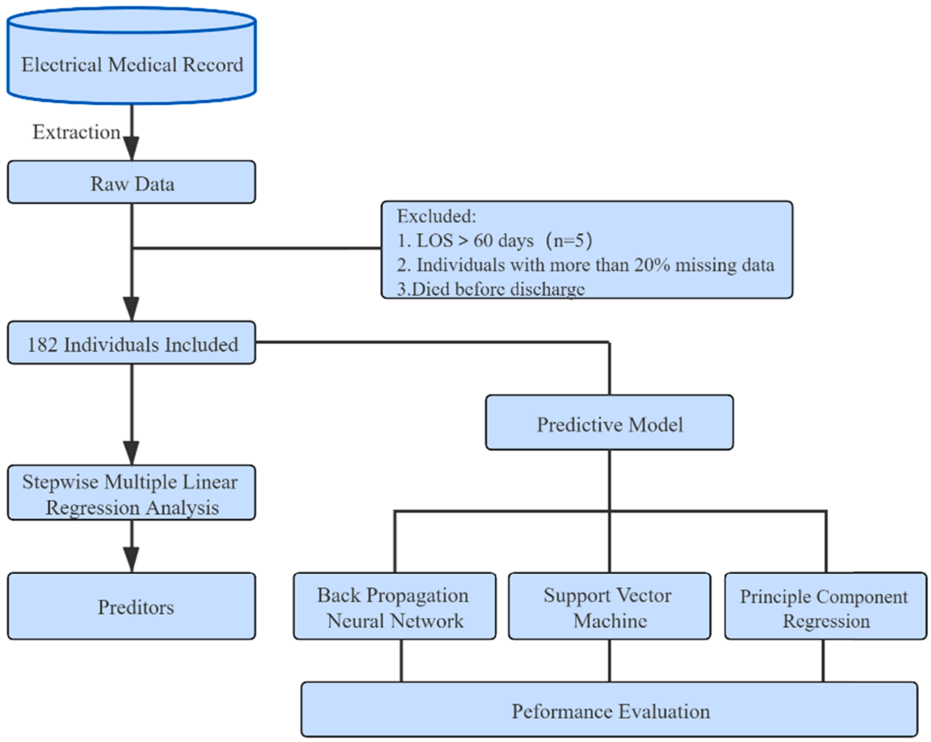

Fig. 2. Architecture of proposed model.

- Number of hidden layer neurons

At present, there is no fixed method to select the number of neurons in the hidden layer. The number of nodes in the hidden layer of the network is determined according to experience and experiments. Too many neurons in the hidden layer will lead to too long learning time. Otherwise, the accuracy of the model will be reduced and the ability of the model to identify unlearned samples will be low. Therefore, the above factors should be considered to design the number of hidden layer neurons. The experiment verified that the number of neurons in the hidden layer was finally 5.

- Activation Function

The selection of activation function, hidden layer function, and output layer function has a great influence on the prediction accuracy of the BP neural network. Generally, the node transfer function of the hidden layer adopts tansig function or logsig function, and the node transfer function of output layer adopts tansig function or purelin function. In this study, the tansig function is used in both hidden layer and output layer.

- Learning Rate

There is no rule to determine the learning rate, can only be tried through experiments, to find the optimal value.

- The Number of Iterations

The number of iterations is selected here 100 times, and the network converges when the error reaches the preset goal of 0.001.

The parameters of the SVM are as follows:

- SVM Type

This study is a regression problem, so the SVM type is selected as Epsilon SVR.

- Kernel function

In this study, multiple kernel functions were tried, and finally Radial basis kernel function was selected

- Other parameters

The mesh method was used to find the optimal parameters for c and g parameters, and the value of the loss function of Epsilon SVR was set to 0.01.

3. Results

3.1. Demographic and clinical characteristics of the participants

A total of 182 patients with femoral neck fracture underwent surgery during the study period. The age and sex distribution showed that 69.2% of the patients were female, and the overall age range was 50–97 years. The most frequent surgical type was total hip arthroplasty (46.7%), which had a slightly higher rate than hemiarthroplasty (45.1%). More than half of the patients had comorbidities. Analysis of the LOS in the total population showed that the mean = 16.6 days and a median = 15.00 days.

3.2. Predictors of the LOS

Stepwise linear regression including all significant factors of LOS showed that preoperative calcium level ($p = 0.017$), preoperative lymphocyte percentage ($p = 0.007$), intraoperative bleeding ($p = 0.041$), glucose and sodium chloride injection after surgery ($p = 0.019$), Charlson Comorbidity Index ($p = 0.007$) and BMI ($p = 0.031$) were significant predictors of LOS (Table 2).

Table 2
Stepwise linear regression(predictors of LOS).

Model	predictor	P value
1	Preoperative calcium	0.019*
2	Preoperative calcium	0.007**
	Preoperative lymphocyte percentages	0.013*
3	Preoperative calcium	0.005**
	Preoperative lymphocyte percentages	0.015*
	Intraoperative bleeding	0.026*
4	Preoperative calcium	0.008**
	Preoperative lymphocyte percentages	0.012*
	Intraoperative bleeding	0.024*
	Liquid medicine 3	0.047*
5	Preoperative calcium	0.009**
	Preoperative lymphocyte percentages	0.008**
	Intraoperative bleeding	0.032*
	Liquid medicine 3	0.028*
	Charlson Comorbidity Index	0.04*
6	Preoperative calcium	**0.017***
	Preoperative lymphocyte percentages	**0.007****
	Intraoperative bleeding	**0.041****
	Liquid medicine 3	**0.019***
	Charlson Comorbidity Index	**0.007****
	BMI	**0.031***

*P < 0.05 **P < 0.01.

3.3. BP, SVM and PCR models

Since the data used for prediction be collected from preoperative stay and surgery, we would already know the preoperative LOS by the time we start prediction. Therefore, our models predict the postoperative LOS (PLOS), we can be informed of the total LOS by adding the predicted postoperative LOS and the known preoperative LOS.

As mentioned earlier in the selection of data sets, for those containing missing values, two schemes were used, namely, direct elimination and mean filling. In addition, there were only 15 samples with hospital stays between 20 and 40 days. As the number of samples within this interval is too small and relatively scattered, it has little effect on the prediction model, or even has a negative impact, so only the samples within 20 days are retained. The sample number of the data set after direct elimination was 110. The sample number of the data set after mean filling was 168. All samples were divided into training set and test set, in which the number of training set samples accounted for 80% of the total samples, and the remaining samples were used as test set to test the prediction performance of samples. To make the results more convincing, the samples used in the training set and test set of the model were randomly selected from the total data set.

Each model was tested 10 times in the test set, and the optimal proportion of the sample with prediction error less than 2 days and 3 days. Because the algorithm of the model using the training set of sample is randomly selected from the total sample, so the prediction results will produce certain randomness. As mentioned earlier, we have two samples: one by direct elimination and the other by average padding, and let's call these two sample sets D-PLOS and A-PLOS. Fig. 3a and b respectively represent the D-PLOS and A-PLOS prediction model, and the optimal proportion of the samples whose errors are less than 2 and 3 days in the total test set. It can be seen from the figure that PCR model shows better performance compared with the other two models. The proportion of samples with errors less than 2 and 3 days is the highest, and the proportion of samples with errors less than 2 days reaches 90.91%. Table 3 shows the Mean Absolute Error (MAE) of 10 tests, the standard deviation (SD), and the optimal MAE of 10 times predicted by three algorithms. In both d-LOS and P-LOS sample sets, MAE and the optimal MAE predicted by PCR prediction model are both the smallest. Compared with A-PLOS, D-LOS has better prediction performance and higher accuracy. Observing SD, it can be found that the stability of SVM model is the best, and that of BP model is the worst.

4. Discussion

Femoral neck fracture is one of the most debilitating conditions, especially in older adults. LOS is an important indicator reflecting hospital productivity. Hospital overcrowding is associated with poor use of resources that may undermine patient safety, whereas underutilization may be associated with a loss of potential revenue. According to the estimation of relevant departments, reducing LOS could provide financial savings hundreds of millions of pounds to NHS hospitals, with a mean higher LOS associated with 4.137 to 4.32 percentage points lower productivity because it reduces the capacity for treating new referrals [12]. On the other hand, prolonged LOS is not just an inefficient use of medical resources but also brings the risk of related complications such as pressure sores, deep vein thrombosis and urinary infections. For every additional day of hospitalization, the chance of complications increases by 5% [5]. Reducing LOS decreased catheter-associated urinary tract infections and the 30-day readmission rate and was accompanied by financial savings for patients [13]. Predicting LOS could also provide an ancillary benefit in which the care of patients with early detection could be optimized preoperatively and the related risk factors could be reduced to minimize the risk of long-term LOS, which could reduce

Table.3
MAE (days), SD and Optimal MAE (days)

Algorithm	Indicator					
	MAE		SD		Optimal MAE	
	D-PLOS	A-PLOS	D-PLOS	A-PLOS	D-PLOS	A-PLOS
BP	3.054	3.6621	0.7194	0.3614	2.2878	3.2710
SVM	2.5824	3.3635	0.1142	0.1265	2.4256	3.2328
PCR	2.1489	3.2075	0.4018	0.3148	1.5254	2.6167

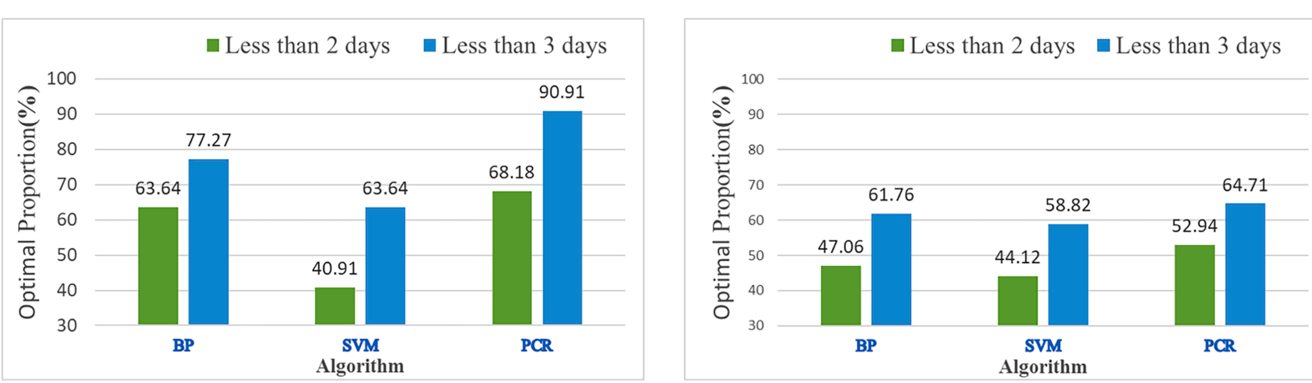

Fig. 3. . Proportion of absolute error within 2 days and 3 days.

patient costs and complication risks [14].

We could not shorten patients' LOS excessively for economic and efficiency gain. Reducing the LOS too aggressively often leads to a lack of medical care. It may deprive patients of optimal care and requisite rehabilitation for functional recovery. A sample cohort study that included 4,213 patients with hip fracture revealed that an LOS of 10 days or less for patients receiving surgery is associated with increased death risk within 1 year of discharge [15]. In Canada, hospitalization after hip fracture surgery shortened evidently in the early 2000 s to improve efficiency. This triggers concern about worsening the quality of medical care resulting from shorter LOS in some care settings [16]. In Sweden, the population aged more than 50 years increased, but hospital beds decreased at the same time. The study showed that shorter LOS after hip fracture was associated with an increased risk of death within 30 days of hospital discharge (when confined to patients with an LOS 10 days throughout the study period). This may be because shorter LOS results in patients' exposure to fewer care providers with enough professional skills in the early postoperative period [17]. In response to the rapidly growing demand for hospital beds, health departments in some countries have adopted incentives or penalties and even formulated policies to reduce LOS, which may lead to some patients who did not reach discharge criteria being discharged early. For these reasons, the hospital can neither shorten the LOS excessively nor be too conservative in increasing the LOS, resulting in wasted resources and medical insurance. Hospital administrators need to minimize the length of stay while meeting strict discharge criteria and reduce the risk of patient-related complications to achieve a win–win situation.

There are substantial differences in the LOS of patients with femoral neck fracture. This sort of discrepancy is determined by both patient's intrinsic related factors and hospital management [18]. We found significant variations in discharge destinations worldwide. Patients discharged to inpatient rehabilitation facilities, nursing homes, and the community in China are different from those in other countries [19]. Discharging to home is the most common choice in China, mainly because of the scarcity of rehabilitation agencies and because discharge marks the end of the main treatment. For this reason, the average LOS (16.6 ± 6.9) for patients with femoral neck fracture was longer in our study, and the main factors affecting LOS were patient related. In China, patients are still prioritizing tertiary hospitals because people believe that large general hospitals provide better quality medical resources. This phenomenon made managing the availability of inpatient beds and coordinating with the volume of surgery a daunting task, especially when the actual hospital LOS was different from the expectations. Tianjin Medical University General Hospital is the largest teaching hospital in Tianjin. A significant proportion of patients have a wide range of comorbidities, and the mean age is older. These factors bring great difficulty in predicting the LOS and compensate for the shortcomings of single-center research to a certain extent.

We collected related anamnesis on admission, laboratory tests before the operation, and intraoperative data of 182 patients by reviewing their electronic records. In this study, preoperative calcium level, intraoperative glucose and sodium chloride injection, and preoperative lymphocyte percentages were predictors associated with postoperative LOS. BMI, CCI, preoperative calcium level, glucose and sodium chloride injection, preoperative lymphocyte percentages, and intraoperative bleeding were predictors associated with LOS, as demonstrated by stepwise regression. Among these factors, preoperative calcium level and BMI are negative factors. The main results are consistent with previous studies. In the present study, we found that excessive volumes of intravenous glucose and sodium chloride infusion after surgery were related to prolonged LOS. Glucose and sodium chloride infusion is used as a source of energy after surgery, and its use is mainly decided by the surgeon's subjective judgment and multiple other factors. For femoral neck fracture patients, fluid overload is harmful, potentially causing gastric dysmotility and poor wound healing [20]. These results suggested that early intake of oral fluids and solids to support energy and protein supply is necessary for patients after surgery [21]. Surgeons should be more aware of fluid therapy strategies after surgery to shorten unnecessary hospital stays. We established three machine learning algorithms. From the research results, the PCR model had the best prediction results. In the PCR modeling process, the total data set had 38 labels, which may be correlated with each other, and there is a multicollinearity problem. Through principal component analysis of these original data, the data label was reduced from the original 38 items to 6 items, which greatly reduced the number of predictive variables, and these 6 items covered most of the information of all indicators. In the process of modeling, if there are too many variables, the establishment of the rule of search will be impeded. In this way, dimensionality reduction simplifies the analysis process and improves the accuracy of the results. Furthermore, we followed the IJMEDI checklist to self assess the quality of the work and avoid bias [22]. The IJMEDI checklist is available on the Zenodo repository (https://zenodo.org/record/4835800#.YRlG29-xU2w).

"The Decision Is More Important Than the Incision." This phrase illustrates the importance of medical decisions [23]. ML allows algorithms to learn from experience without clearly being programmed. The application of machine learning enables us to maximize the use of demographic data, selecting clinical strategies early to better adjust the length of stay. Machine learning algorithms and the combination of many kinds of data streams can guide all health system administrators, medical personnel and patients and their families. It could provide more systematic medical services for patients, offering more effective information, including funding and nursing for social medical insurance. Precise prediction of LOS allows hospital administrators to plan the number of beds and members of staff required to meet the demand for care more appropriately, identifying patients who may need longer LOS for direct care quality improvement and utilize of scarce resources more reasonably. Patients could be informed of discharge plans in advance. The therapeutic schedule could be more suited for the individual patient as opposed to solely relying on population-based studies, fitting the growing trend toward personalized medicine. ML has great potential as a supplementary source of medical information that can help guide the process of surgical and medical decision making [24]. This study still has a few limitations that should be mentioned. The sample size was not large enough, which restricted machine learning from achieving better results. We plan to include more kinds of diseases to improve generalizability in the future. The findings suggested the feasibility of using machine learning to predict LOS for patients with femoral neck fracture, and this field needs more investigation.

In this study, we offering hospital administrators an effective tool for predicting the PLOS. We identified several predictors of length of stay, enabling physicians to make better clinical decisions to accelerate patient recovery and discharge which conform to the ERAS (Enhanced Recovery After Surgery) concept. We measure the model's performance by MAE, which is straightforward to doctors than other evaluation metrics. Based on our study, we believe that the models we proposed could aid clinicians in overcoming several challenges to timely discharges. We predicted an exact value directly, outputting a specific value instead of classifying LOS into strata as mentioned in the above studies. Despite the success of neural networks on image, text, and speech data, they show weak outcomes in our tabular data. PCR is the most accurate and efficient algorithm among the several models we established.

5. Conclusion

Predicting LOS fits with the development trend of personalized medicine. We used traditional regression to identify predictors of LOS. Excessive intravenous glucose and sodium chloride infusion after surgery, preoperative hypocalcemia, preoperative high percentages of lymphocytes, excessive intraoperative bleeding, lower BMI and higher CCI scores are related to prolonged LOS. As a flexible technique designed to learn and generalize from "big data", machine learning is a well-suited powerful tool to help hospital administrators make clinical decisions early to avoid related risks and rationally allocate medical resources. In the present study, we successfully constructed multiple algorithms to predict PLOS for patients with femoral neck fracture by using preoperative and intraoperative data. PCR showed optimal results in predicting PLOS. Future work could be extended to include other diseases.

6. Summary table

What was already known on the topic:

- Femoral neck fracture is a frequent cause of hospitalization in the orthopedic department. Additionally, LOS has been used as a common surrogate marker for hospital cost and quality of care provided.
- Constructing an accurate algorithm is difficult. Traditional statistical methods are standard when developing prediction models but are not as accurate as machine learning.

What this study added to our knowledge:

- Stepwise linear regression could identify predictors of LOS.
- Machine learning could predict PLOS accurately, and the principal component regression shows better performance.

CRediT authorship contribution statement

Hao Zhong: Conceptualization, Writing – original draft. **Bingpu Wang:** Writing – original draft, Methodology. **Dawei Wang:** Conceptualization. **Zirui Liu:** Methodology, Investigation. **Cong Xing:** Formal analysis. **Yu Wu:** Formal analysis. **Qiang Gao:** Formal analysis. **Shibo Zhu:** Investigation. **Haodong Qu:** Writing – original draft. **Zeyu Jia:** Methodology, Investigation. **Zhigang Qu**: Writing – review & editing. **Guangzhi Ning:** Writing – review & editing. **Shiqing Feng:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2021.104572.

References

[1] C. Ji, Y. Zhu, S. Liu, et al., Incidence and risk of surgical site infection after adult femoral neck fractures treated by surgery: A retrospective case-control study, Medicine (Baltimore) 98 (11) (2019) e14882.
[2] C.-L. Cheung, S.B. Ang, M. Chadha, E.-L. Chow, Y.-S. Chung, F.L. Hew, U. Jaisamrarn, H. Ng, Y. Takeuchi, C.-H. Wu, W. Xia, J. Yu, S. Fujiwara, An updated hip fracture projection in Asia: The Asian Federation of Osteoporosis Societies study, Osteoporos. Sarcopenia 4 (1) (2018) 16–21.
[3] N. Veronese, S. Maggi, Epidemiology and social costs of hip fracture, Injury 49 (8) (2018) 1458–1460.
[4] P.P. Goodney, T.A. Stukel, F.L. Lucas, et al., Hospital volume, length of stay, and readmission rates in high-risk surgery, Ann. Surg. 238 (2) (2003) 161–167.
[5] P.J. Mathew, F. Jehan, N. Kulvatunyou, M. Khan, T. O'Keeffe, A. Tang, L. Gries, M. Hamidi, E.-R. Zakaria, B. Joseph, The burden of excess length of stay in trauma patients, Am. J. Surg. 216 (5) (2018) 881–885.
[6] K.C. Safavi, T. Khaniyev, M. Copenhaver, et al., Development and validation of a machine learning model to aid discharge processes for inpatient surgical care, JAMA Netw. Open 2 (12) (2019) e1917221.
[7] A.E. Garcia, J.V. Bonnaig, Z.T. Yoneda, et al., Patient variables which may predict length of stay and hospital costs in elderly patients with hip fracture, J. Orthopaedic Trauma 26 (11) (2012) 620–623.
[8] J.E. Clague, E. Craddock, G. Andrew, M.A. Horan, N. Pendleton, Predictors of outcome following hip fracture. Admission time predicts length of stay and in-hospital mortality, Injury 33 (1) (2002) 1–6.
[9] M.M. Churpek, T.C. Yuen, C. Winslow, D.O. Meltzer, M.W. Kattan, D.P. Edelson, Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards, Crit. Care Med. 44 (2) (2016) 368–374.
[10] A.M. Darcy, A.K. Louie, L.W. Roberts, Machine learning and the profession of medicine, JAMA 315 (6) (2016) 551–552.
[11] M. Huber, C. Kurz, R. Leidl, Predicting patient-reported outcomes following hip and knee replacement surgery using supervised machine learning, BMC Med. Inform. Decis Mak. 19 (1) (2019) 3.
[12] M. Ali, R. Salehnejad, M. Mansur, Hospital productivity: The role of efficiency drivers, Int. J. Health Plann. Manage. 34 (2) (2019) 806–823.
[13] P.N. Ramkumar, S.M. Navarro, H.S. Haeberle, J.M. Karnuta, M.A. Mont, J. P. Iannotti, B.M. Patterson, V.E. Krebs, Development and validation of a machine learning algorithm after primary total hip arthroplasty: applications to length of stay and payment models, J. Arthroplasty 34 (4) (2019) 632–637.
[14] R.A. Gabriel, B.S. Sharma, C.N. Doan, X. Jiang, U.H. Schmidt, F. Vaida, A predictive model for determining patients not requiring prolonged hospital length of stay after elective primary total hip arthroplasty, Anesth. Analg. 129 (1) (2019) 43–50.
[15] J. Yoo, J.S. Lee, S. Kim, B.S. Kim, H. Choi, D.Y. Song, W.B. Kim, C.W. Won, Length of hospital stay after hip fracture surgery and 1-year mortality, Osteoporos. Int. 30 (1) (2019) 145–153.
[16] B. Sobolev, P. Guy, K.J. Sheehan, et al., Hospital mortality after hip fracture surgery in relation to length of stay by care delivery factors: A database study, Medicine (Baltimore) 96 (16) (2017) e6683.
[17] P. Nordstrom, Y. Gustafson, K. Michaelsson, et al., Length of hospital stay after hip Fracture and short term risk of death after discharge: a total cohort study in Sweden, BMJ 350 (2015) h696.
[18] T. Richards, A. Glendenning, D. Benson, S. Alexander, S. Thati, The independent patient factors that affect length of stay following hip fractures, Ann. R Coll Surg. Engl. 100 (7) (2018) 556–562.
[19] K.B. Pitzul, W.P. Wodchis, H.J. Kreder, et al., Discharge destination following hip fracture: comparative effectiveness and cost analyses, Archives Osteoporosis 12 (1) (2017) 87.
[20] S.R. Lewis, A.R. Butler, A. Brammar, et al., Perioperative fluid volume optimization following proximal femoral fracture, Cochrane Database System. Rev. 3 (2016). CD003004.
[21] O. Ljungqvist, M. Scott, K.C. Fearon, Enhanced recovery after surgery, JAMA Surg. 152 (3) (2017).
[22] F. Cabitza, A. Campagner, The need to separate the wheat from the chaff in medical informatics: Introducing a comprehensive checklist for the (self)-assessment of medical AI studies, Int. J. Med. Inform. 153 (2021), 104510.
[23] M.D. Moisi, J. Page, S. Gahramanov, R.J. Oskouian, Bullet fragment of the lumbar spine: the decision is more important than the incision, Global Spine J. 5 (6) (2015) 523–526.
[24] J.T. Senders, P.C. Staples, A.V. Karhade, M.M. Zaki, W.B. Gormley, M.L. D. Broekman, T.R. Smith, O. Arnaout, Machine learning and neurosurgical outcome prediction: A systematic review, World Neurosurg. 109 (2018) 476–486. e1.