MEDS Clinical Medicine (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/medsc.2026.070308
ISSN 2616-1907 Vol. 7 Num. 3

Development and validation of a machine learning-based prediction model for surgical site infection after spinal metastasis surgery

Qiang Ren[1], Qichang Li[1,*]

[1] *Department of Bone and Soft Tissue Oncology, Chongqing University Cancer Hospital, Chongqing, 400030, China*

**Corresponding author. trees0030@163.com*

Keywords: Spinal Metastases; Surgical Site Infection; Prediction Model; Machine Learning

Abstract: Surgical site infection (SSI) is a severe postoperative complication in patients with spinal metastases. This study retrospectively enrolled 460 patients with spinal metastases who underwent surgery at a single center. Stepwise variable screening using univariate analysis, LASSO regression, and multivariate Logistic regression identified seven independent risk factors for SSI: operative time, age, ECOG score, diabetes mellitus, open surgery, preoperative chemotherapy, and hypoproteinemia. Six machine learning models were constructed, and 10-fold cross-validation showed that the Logistic regression model achieved the best performance (AUC = 0.906), with an AUC of 0.866 in the test set. This study developed and validated a risk prediction model for SSI following spinal metastasis surgery with good discrimination and clinical interpretability, providing a reference for individualized perioperative infection risk assessment.

1. Introduction

The spine represents the most common site of osseous metastasis from malignant tumors. Approximately 10% of patients with spinal metastases develop metastatic epidural spinal cord compression, manifesting as pain, pathological fractures, and neurological deficits [1]. With advances in comprehensive cancer therapy, patient survival has improved substantially, leading to a growing number of patients with spinal metastases undergoing surgical intervention each year. However, these patients are immunocompromised and undergo extensive surgical procedures, resulting in a markedly higher risk of postoperative complications compared with conventional spinal surgery. Among these, surgical site infection (SSI) is one of the most prevalent and serious complications [2].

Previous systematic reviews have demonstrated that the overall incidence of SSI following surgery for spinal metastases is approximately 6.5%, representing the leading cause of reoperation and accounting for 27.8% of all reoperation indications [2]. The reported incidence varies considerably across studies, ranging from 3.85% to 14.7%. SSI not only prolongs hospital length of stay and increases healthcare costs but also delays the initiation of adjuvant chemoradiotherapy, thereby adversely affecting patient prognosis. Accordingly, the preoperative identification of high-risk patients and the implementation of targeted preventive measures carry important clinical implications.

Numerous studies have reported risk factors associated with SSI after spinal surgery, and the predictive value of factors such as diabetes mellitus, operative time, preoperative radiotherapy, obesity, and malnutrition has been widely recognized [3]. With the widespread application of artificial intelligence technology in clinical research, machine learning models have been increasingly used for predicting postoperative infection risk in spinal surgery, and related studies have confirmed their favorable predictive potential [4]. This study retrospectively collected clinical data from 460 patients with spinal metastases who underwent surgery at a single center. LASSO regression and multivariate Logistic regression were used to identify independent risk factors for postoperative SSI. Multiple machine learning prediction models were constructed and internally validated, and their predictive performances were compared. This study aims to provide a reference for preoperative infection risk assessment in patients with spinal metastases.

2. Materials and Methods

2.1. Study Population and Clinical Data Collection

This study was a single-center retrospective study that enrolled 460 patients with spinal metastases who underwent surgical treatment at our hospital from January 2020 to May 2026. Clinical data were retrospectively collected through the electronic medical record system and categorized as follows: (1) demographic characteristics: age, sex, body mass index (BMI), drinking history; (2) comorbidities: diabetes mellitus, hypoproteinemia; (3) functional status: Eastern Cooperative Oncology Group Performance Status (ECOG) score; (4) tumor-related characteristics: primary tumor type (breast cancer, prostate cancer, gastrointestinal cancer, lung cancer, and others), preoperative chemotherapy history; (5) surgery-related factors: surgical approach (open surgery/minimally invasive surgery), operative time, anesthesia method, intraoperative blood transfusion; (6) preoperative laboratory parameters: white blood cell count (WBC) and hemoglobin (Hb) within 1 week before surgery.

2.2. Outcome Measures

The primary outcome was surgical site infection (SSI). The diagnosis was based on the Chinese National Health Commission health industry standard WS/T 861-2025, largely consistent with the CDC NHSN criteria. Postoperative day 30 was set as the observation window. Superficial incisional, deep incisional, and organ/space SSI were combined for unified diagnosis. SSI was confirmed if any of the following criteria were met within 30 days postoperatively: (1) purulent discharge from the surgical incision; (2) purulent exudate obtained from deep tissues via needle aspiration, wound opening, or reoperation; (3) pathogenic bacteria isolated from bacterial culture; (4) unexplained persistent fever (≥38°C) with local inflammatory signs responsive to anti-infective therapy, excluding infection at other sites; (5) characteristic infectious changes on CT or MRI, including abscess, inflammatory exudate, or inflammatory bone destruction.

2.3. Statistical Analysis

Statistical analysis and modeling were performed using Python 3.10. Continuous variables are presented as mean ± standard deviation ($\bar{x}\pm s$) for normally distributed data, with intergroup comparisons using the independent Student's t-test. Non-normally distributed data are presented as median (interquartile range) [M (IQR)], with comparisons using the Mann-Whitney U test. Categorical variables are presented as number (percentage) [n (%)], with comparisons using the χ^2 test. A P-value < 0.05 was considered statistically significant.

2.4. Variable Selection

Variables were screened using a three-step hierarchical approach. Step 1: Univariate Logistic regression was used to identify potential risk factors, and variables with $P < 0.20$ were included in the candidate variable set. Step 2: Multicollinearity was assessed using the variance inflation factor (VIF), with VIF > 10 indicating significant collinearity. Subsequently, LASSO regression with 10-fold cross-validation was applied to select key predictive variables with non-zero coefficients using the λ.1se optimal regularization parameter. Step 3: The selected variables were entered into multivariate Logistic regression, and variables with $P < 0.05$ were identified as independent risk factors.

2.5. Machine Learning Model Construction and Evaluation

The entire study cohort was divided into a training set and a test set at an 8:2 ratio using stratified random sampling based on postoperative SSI status. Six machine learning algorithms were used: Logistic Regression (LR), Gradient Boosting Machine (GBM), XGBoost, Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVM). All models were trained using the same set of predictors. Ten-fold stratified cross-validation was performed within the training set, with evaluation metrics including AUC, accuracy, sensitivity, specificity, F1-score, PPV, and NPV. The model with the highest mean cross-validation AUC was selected as optimal.

2.6. Model Validation

The predictive performance of the optimal model was further validated on the independent test set. The discriminative ability was evaluated using AUC and its 95% confidence interval (CI). Additional metrics including accuracy, sensitivity, specificity, F1-score, PPV, and NPV were also calculated.

2.7. SHAP Interpretability Analysis

SHapley Additive exPlanations (SHAP) was applied to conduct interpretability analysis of the optimal model. A SHAP beeswarm plot was generated to visually demonstrate the direction and magnitude of each predictor's contribution to the model output, further elucidating the underlying mechanisms by which each risk factor influences the risk of postoperative SSI.

3. Results

3.1. Baseline Patient Characteristics

A total of 460 patients with spinal metastases who underwent surgery were included. Among them, 32 patients developed postoperative SSI, with an infection rate of 7.0%; 428 patients did not develop infection, accounting for 93.0%. Intergroup comparisons are shown in Table 1.

Table 1. Comparison of baseline characteristics between the two groups

Variable	**Non-infection group (n=428)**	**Infection group (n=32)**	**P-value**
Operative time (min)	124.00 (86.75, 164.00)	167.00 (121.00, 197.25)	<0.001
Age (years)	61.00 (56.00, 67.00)	69.00 (64.75, 75.50)	<0.001
BMI (kg/m²)	22.30 (20.10, 23.80)	22.10 (20.38,	0.977

		23.62)	
Preoperative WBC count ($\times 10^9$/L)	6.00 (4.90, 7.00)	6.15 (4.65, 6.95)	0.915
Preoperative hemoglobin (g/L)	112.75 ±15.83	112.03 ±16.10	0.805
ECOG score			<0.001
2	230 (53.7%)	8 (25.0%)	
3	132 (30.8%)	11 (34.4%)	
4	66 (15.4%)	13 (40.6%)	
Diabetes mellitus			<0.001
No	328 (76.6%)	12 (37.5%)	
Yes	100 (23.4%)	20 (62.5%)	
Surgical approach			0.002
Open surgery	233 (54.4%)	28 (87.5%)	
Minimally invasive surgery	195 (45.6%)	4 (12.5%)	
Preoperative chemotherapy			0.002
No	213 (49.8%)	7 (21.9%)	
Yes	215 (50.2%)	25 (78.1%)	
Preoperative hypoproteinemia			<0.001
No	254 (59.3%)	6 (18.8%)	
Yes	174 (40.7%)	26 (81.2%)	
Sex			0.126
Female	207 (48.4%)	11 (34.4%)	
Male	221 (51.6%)	21 (65.6%)	
Smoking history			0.105
No	304 (71.0%)	27 (84.4%)	
Yes	124 (29.0%)	5 (15.6%)	
Coronary heart disease			0.831
No	367 (85.7%)	27 (84.4%)	
Yes	61 (14.3%)	5 (15.6%)	
Surgical level			0.761
Thoracic	153 (35.7%)	14 (43.8%)	
Lumbar	179 (41.8%)	13 (40.6%)	
Cervical	40 (9.3%)	2 (6.2%)	
Sacral	56 (13.1%)	3 (9.4%)	
Hypertension			0.886
No	233 (54.4%)	17 (53.1%)	
Yes	195 (45.6%)	15 (46.9%)	
Primary tumor type			0.853
Lung cancer	119 (27.8%)	10 (31.2%)	
Prostate cancer	85 (19.9%)	4 (12.5%)	
Breast cancer	75 (17.5%)	6 (18.8%)	
Gastrointestinal cancer	74 (17.3%)	5 (15.6%)	
Others	75 (17.5%)	7 (21.9%)	

3.2. Univariate Logistic Regression Analysis

Univariate Logistic regression analysis showed that seven variables were significantly associated with postoperative SSI (all $P < 0.05$), including operative time (OR = 1.015), age (OR = 1.125), ECOG score (OR = 2.379), diabetes mellitus (OR = 5.467), surgical approach (OR = 5.858), preoperative chemotherapy (OR = 3.538), and preoperative hypoproteinemia (OR = 6.326). According to the $P < 0.20$ criterion, these seven variables together with sex ($P = 0.123$) and smoking history ($P = 0.087$) were included in the candidate variable set. A total of nine variables were entered into subsequent modeling.

3.3. Variable Selection by LASSO Regression

Ten-fold cross-validation for LASSO regression revealed that $\lambda.min = 1.260$ and $\lambda.1se = 12.751$. At the $\lambda.min$ threshold, the regression coefficients of all nine candidate variables were non-zero and all were retained. At the $\lambda.1se$ threshold, the coefficients of sex and smoking history were shrunk to zero and thus excluded, leaving seven variables for further analysis.

3.4. Multivariate Logistic Regression Analysis

The seven variables selected by LASSO regression were entered into a multivariate Logistic regression model. The results showed that all seven variables were independent risk factors for postoperative SSI (all $P < 0.05$). Detailed statistical results are presented in Table 2.

Table 2. Multivariate Logistic regression analysis: independent risk factors for postoperative SSI

Variable	OR	95% CI	P-value
Operative time (min)	1.020	1.009–1.031	<0.001
Age (years)	1.156	1.089–1.228	<0.001
ECOG score	3.073	1.625–5.809	<0.001
Diabetes mellitus (yes)	7.704	2.777–21.371	<0.001
Surgical approach (open)	4.245	1.221–14.759	0.023
Preoperative chemotherapy (yes)	5.452	1.633–18.200	0.006
Preoperative hypoproteinemia (yes)	4.571	1.499–13.939	0.008

3.5. Performance Comparison of Machine Learning Models

Six machine learning prediction models were constructed based on the seven independent risk factors. The results of 10-fold cross-validation are presented in Table 3. The Logistic regression model achieved the highest mean AUC (0.906 ± 0.096), as shown in Figure 1, followed by random forest, GBM, XGBoost, SVM, and decision tree. XGBoost had the highest accuracy (0.913 ± 0.064), while Logistic regression had the highest sensitivity (0.783 ± 0.269). The negative predictive values (NPV) of all models were relatively high, ranging from 0.968 to 0.983. Considering both predictive performance and clinical interpretability, Logistic regression was finally selected as the optimal prediction model.

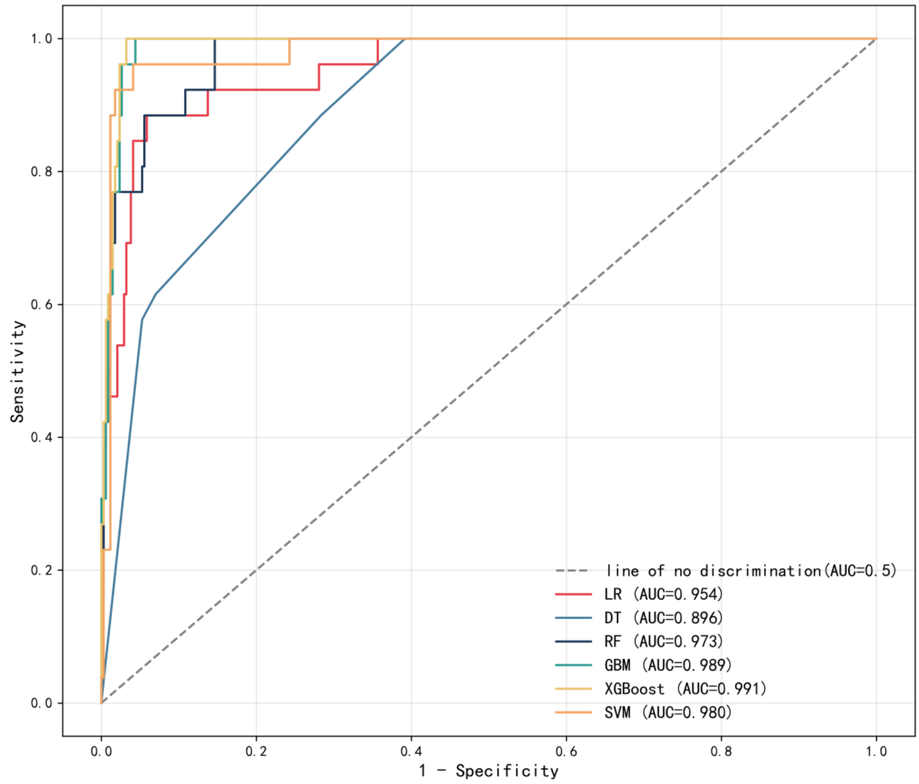

Figure 1. Comparison of ROC curves of machine learning models in the training set

Table 3. Performance comparison of machine learning models in 10-fold cross-validation on the training set

Model	AUC (mean ± SD)	Accuracy (mean ± SD)	Sensitivity (mean ± SD)	Specificity (mean ± SD)	F1-score (mean ± SD)	PPV (mean ± SD)	NPV (mean ± SD)
LR	0.906 ± 0.096	0.872 ± 0.064	0.783 ± 0.269	0.877 ± 0.060	0.477 ± 0.178	0.347 ± 0.137	0.983 ± 0.023
RF	0.862 ± 0.120	0.908 ± 0.059	0.617 ± 0.269	0.930 ± 0.048	0.505 ± 0.244	0.448 ± 0.251	0.969 ± 0.024
GBM	0.845 ± 0.148	0.886 ± 0.056	0.633 ± 0.340	0.903 ± 0.040	0.448 ± 0.262	0.349 ± 0.215	0.971 ± 0.026
XGBoost	0.844 ± 0.159	0.913 ± 0.064	0.633 ± 0.340	0.933 ± 0.047	0.529 ± 0.314	0.461 ± 0.299	0.972 ± 0.026
SVM	0.831 ± 0.166	0.872 ± 0.072	0.633 ± 0.340	0.888 ± 0.068	0.404 ± 0.232	0.308 ± 0.198	0.971 ± 0.027
DT	0.709 ± 0.184	0.652 ± 0.113	0.667 ± 0.325	0.649 ± 0.130	0.205 ± 0.108	0.123 ± 0.067	0.968 ± 0.029

3.6. Validation of the Optimal Model

On the independent test set, the AUC of the Logistic regression model was 0.866, which was lower than the mean AUC of 10-fold cross-validation on the training set (0.906), suggesting a certain degree of overfitting.

3.7. SHAP Interpretability Analysis

SHAP analysis revealed that operative time, age, and diabetes mellitus were the three features with the greatest contribution to model prediction (Figure 2). Longer operative time, older age, and diabetes mellitus were associated with a higher predicted probability of SSI. ECOG score, preoperative hypoproteinemia, preoperative chemotherapy, and surgical approach also made substantial contributions, and the direction of effect for each variable was consistent with the results of multivariate Logistic regression.

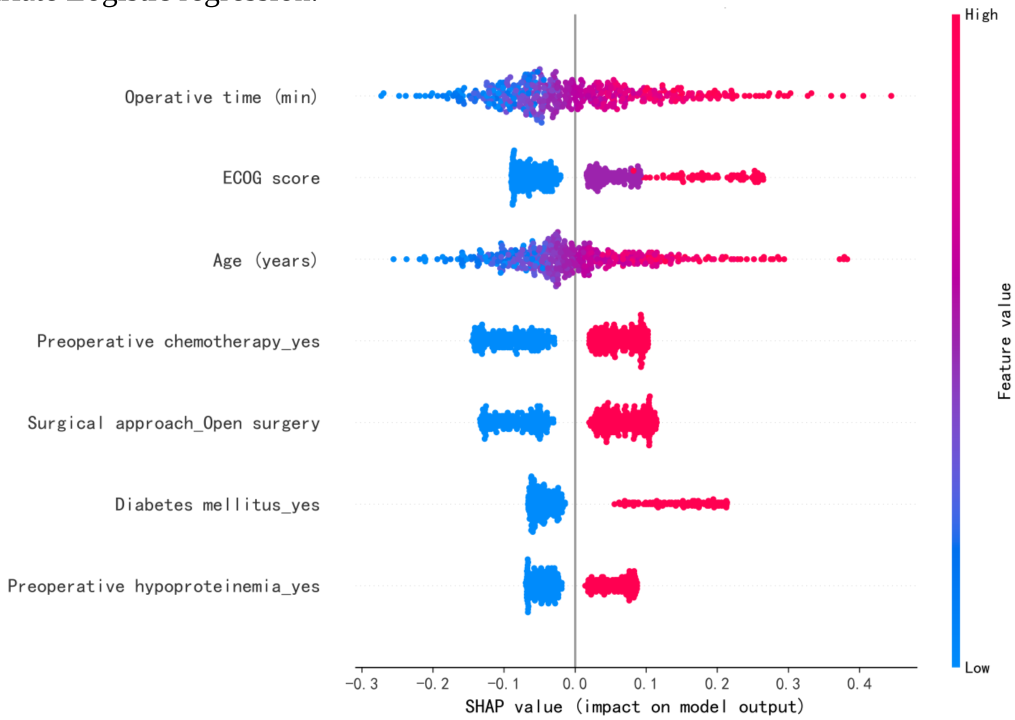

Figure 2. SHAP beeswarm plot for the LR model predicting postoperative SSI in spinal metastases

4. Discussion

Based on clinical data from 460 patients, this study showed that the incidence of SSI within 30 days postoperatively was 7.0%, which is generally consistent with previous literature. A single-center study including 297 patients reported an SSI incidence of 5.1% [5], while a systematic review of 19 studies involving 2088 patients showed a pooled incidence of 6.5% [2], with some studies reporting rates as high as 11.2% [6]. The variation across studies may be related to diagnostic criteria, follow-up duration, and patient baseline characteristics. Importantly, SSI is the leading cause of reoperation in this population [2], and given that surgery for spinal metastases is primarily aimed at palliation [7], effective prevention and control of postoperative SSI is of particularly important clinical significance.

Accurate risk assessment of SSI after spinal metastasis surgery has long been a challenge in clinical decision-making, as current clinical judgment mostly relies on postoperative symptoms and inflammatory markers. In recent years, machine learning techniques have shown promising applications in predicting perioperative complications, with previous studies achieving favorable discriminative performance [4]. In this study, six machine learning algorithms were compared, and the Logistic regression model achieved the highest mean AUC (0.906) in 10-fold cross-validation, suggesting that traditional statistical regression methods still have distinct advantages in data structures with a moderate number of variables and predominantly linear effect relationships.

Multivariate Logistic regression analysis identified seven independent risk factors for SSI. Diabetes mellitus was one of the strongest risk factors, consistent with previous studies, as a hyperglycemic environment impairs neutrophil function and microcirculatory perfusion, thereby delaying wound healing [5]. Operative time was also confirmed as an independent risk factor, with longer operative time prolonging incision exposure and increasing intraoperative blood loss [8]. The development of minimally invasive techniques has provided new avenues for shortening operative time and reducing tissue damage, with previous studies confirming that minimally invasive surgery can reduce postoperative infection rates [9]. Preoperative hypoproteinemia and ECOG score reflect nutritional reserve and functional status, respectively, both of which are closely associated with impaired immune function. Preoperative chemotherapy increases infection risk by inducing myelosuppression, while open surgery carries a higher risk than minimally invasive surgery. Age, as an uncontrollable host factor, is also associated with SSI due to declined immune function and impaired tissue repair capacity.

This study has certain limitations. First, this is a single-center retrospective study, and the generalizability of the conclusions is limited to some extent. Second, the AUC on the test set was lower than that of 10-fold cross-validation, suggesting mild overfitting. In addition, this study only included routine preoperative and perioperative clinical indicators, without covering potential factors such as dynamic changes in nutritional status and inflammatory markers. Future studies could further optimize predictive performance through multicenter, large-sample, prospective cohort studies with external validation.

Acknowledgement

Funding: This work was supported by the 2025 Annual Technological Innovation Project of Shapingba District, Chongqing (Grant No. 2025116).

Institutional Review Board Statement: This research was approved by the Medical Research Ethics Committee of Chongqing University Cancer Hospital.

References

[1] Fisher CG, DiPaola CP, Ryken TC, et al. A novel classification system for spinal instability in neoplastic disease: an evidence-based approach and expert consensus from the Spine Oncology Study Group. Spine (Phila Pa 1976). 2010;35(22):E1221-E1229.

[2] Tarawneh AM, Pasku D, Quraishi NA. Surgical complications and re-operation rates in spinal metastases surgery: a systematic review. Eur Spine J. 2021;30(10):2791-2799.

[3] Rosenke SL, Kisekka M, Darweesh H, et al. Risk factors for surgical site infections after spinal surgery: a systematic review and meta-analysis. Eur Spine J. 2026;35(7):4181-4197.

[4] Cui Y, Shi X, Wang Q, et al. Artificial intelligence-based prediction model for surgical site infection in metastatic spinal disease: a multicenter development and validation study. Int J Surg. 2025;111(10):6867-6884.

[5] Sebaaly A, Shedid D, Boubez G, et al. Surgical site infection in spinal metastasis: incidence and risk factors. Spine J. 2018;18(8):1382-1387.

[6] Atkinson RA, Davies B, Jones A, et al. Survival of patients undergoing surgery for metastatic spinal tumours and the impact of surgical site infection. J Hosp Infect. 2016;94(1):80-85.

[7] Kurisunkal V, Gulia A, Gupta S. Principles of Management of Spine Metastasis. Indian J Orthop. 2020;54(1):181-193.

[8] Igoumenou VG, Mavrogenis AF, Angelini A, et al. Complications of spine surgery for metastasis. Eur J Orthop Surg Traumatol. 2020;30(1):37-56.

[9] McCabe FJ, Jadaan MM, Byrne F, et al. Spinal metastasis: The rise of minimally invasive surgery. Surgeon. 2022;20(5):328-333.