Construction of a Risk Prediction Model for Diabetic Foot Ulcer Recurrence: Based on Logistic Regression, Support Vector Machine, and BP Neural Network Models (Postprint)

Authors: Zhang Juan, Li Haifen, Li Xiaoman, Yao Miao, Ma Huizhen, Ma Qiang *

Date: 2023-07-25T00:00:00+00:00

Abstract

Background: The rates of first and subsequent recurrences of diabetic foot ulcers (DFUs) are increasing globally, with early recurrence risk being higher than long-term risk. There are numerous risk factors for DFU recurrence, and systematic screening is currently lacking. Therefore, it is necessary to explore the risk factors for DFU recurrence to enable early identification of high-risk populations. Objective: To investigate the predictive value of Logistic regression, Support Vector Machine (SVM), and Backpropagation Neural Network (BPNN) models in assessing the risk of diabetic foot ulcer (DFU) recurrence. Methods: A total of 390 DFU patients who visited the Department of Burn, Plastic and Cosmetic Surgery and the Wound and Stoma Clinic of the General Hospital of Ningxia Medical University between January 2020 and December 2022 were selected as study subjects for model development. Patients were divided into a recurrence group (116 cases, 29.7%) and a non-recurrence group (274 cases, 70.3%) based on whether DFU recurred within 1 year after discharge. General data of patients in both groups were collected, including sociodemographic characteristics, medical history assessment, and clinical case data, and compared. The Diabetic Foot Self-Management Behavior Scale (DFSBS) was used to assess patients' diabetic foot self-management behaviors, and the Chronic Disease Risk Perception Questionnaire was used to evaluate patients' risk perception levels regarding DFUs. Multivariate Logistic regression analysis was employed to explore the influencing factors of DFU recurrence within 1 year after discharge. Patients were divided into training and test sets at a 70:30 ratio, and Logistic regression variable screening strategies were applied to develop Logistic regression, SVM, and BPNN models, respectively. Receiver Operating Characteristic (ROC) curves for each model predicting

DFU recurrence risk were plotted, and predictive performance was compared using precision, recall, accuracy, F1-score, and Area Under the ROC Curve (AUC). Results: Comparisons between the two groups of DFU patients revealed statistically significant differences in age, BMI, time to medical consultation after onset, length of hospital stay, smoking history, foot ulcer grade, history of amputation of affected toes, ulcer location on the plantar surface, living alone, presence of walking impairment in the foot, ulcer caused by trauma, ankle-brachial index, diabetic peripheral neuropathy, lower extremity atherosclerosis, multidrug-resistant bacterial infection, glycated hemoglobin, albumin, foot care behavior, and risk perception level ($P<0.05$). Multivariate Logistic regression analysis showed that length of hospital stay [OR=0.678, 95%CI (0.557, 0.826), $P<0.001$], ulcer location on the plantar surface [OR=0.078, 95%CI (0.011, 0.541), P=0.010], living alone [OR=5.689, 95%CI (2.583, 10.726), $P<0.001$], presence of walking impairment in the foot [OR=3.364, 95%CI (2.742, 5.638), $P<0.001$], diabetic peripheral neuropathy [OR=3.089, 95%CI (1.156, 4.585), P=0.025], lower extremity atherosclerosis [OR=4.033, 95%CI (3.688, 9.060), $P<0.001$], multidrug-resistant bacterial infection [OR=3.241, 95%CI (1.728, 7.361), $P<0.001$], glycated hemoglobin [OR=0.209, 95%CI (0.065, 0.669), P=0.008], albumin [OR=2.038, 95%CI (1.515, 2.741), $P<0.001$], foot care behavior [OR=1.965, 95%CI (0.874, 3.208), P=0.014], and risk perception level [OR=0.261, 95%CI (0.197, 0.825), P=0.002] were influencing factors for DFU recurrence within 1 year ($P<0.05$). The accuracy rates of Logistic regression, SVM, and BPNN models in predicting DFU recurrence risk in the test set were 82.03%, 94.87%, and 87.17%, respectively, with AUC values of 0.843, 0.930, and 0.800, respectively. Comparison of ROC curve AUCs among the Logistic regression, SVM, and BPNN models for predicting DFU recurrence risk showed a statistically significant difference (Z=8.826, $P<0.05$). The ROC curve AUC of the SVM model was higher than that of the Logistic regression and BPNN models (Z=5.672, P=0.014; Z=2.767, $P<0.001$). Conclusion: The SVM model demonstrated favorable performance in terms of accuracy, sensitivity, specificity, AUC, and other metrics for predicting DFU recurrence risk within 1 year after discharge, making it the relatively optimal model. Further promotion and application are recommended to validate the predictive efficacy of the model.

Full Text

Preamble

Construction of a Recurrence Risk Prediction Model for Diabetic Foot Ulcer Based on Logistic Regression, Support Vector Machine, and BP Neural Network Models

ZHANG Juan, LI Haifen, LI Xiaoman, YAO Miao, MA Huizhen, MA Qiang*

Department of Burn and Plastic Surgery, General Hospital of Ningxia Medical University, Yinchuan 750001, China

*Corresponding author: MA Qiang, Associate Chief Physician; E-mail: 10901969@qq.com

Abstract

Background: The global rates of first and multiple recurrences of diabetic foot ulcers (DFUs) are increasing annually, with early recurrence risk higher than long-term risk. Numerous risk factors contribute to DFU recurrence, yet systematic screening remains lacking. Therefore, exploring risk factors for DFU recurrence is essential for early identification of high-risk populations.

Objective: To investigate the predictive value of Logistic regression (LR), support vector machine (SVM), and BP neural network (BPNN) models in assessing DFU recurrence risk.

Methods: A total of 390 DFU patients treated at the Department of Burn and Plastic Surgery and Wound Ostomy Clinic of Ningxia Medical University General Hospital between January 2020 and December 2022 were enrolled as the study cohort for model development. Patients were divided into a recurrence group (n=116, 29.7%) and a non-recurrence group (n=274, 70.3%) based on DFU recurrence within one year post-discharge. General data including sociodemographic characteristics, medical history assessment, and clinical case information were collected and compared between groups. The Diabetes Foot Self-care Behavior Scale (DFSBS) was used to evaluate patients' foot self-management behaviors, while the Chronic Disease Risk Perception Questionnaire assessed their DFU risk perception levels. Multivariable Logistic regression analysis was employed to identify factors influencing DFU recurrence within one year post-discharge. Patients were randomly split into training and test sets at a 7:3 ratio. Using a Logistic regression variable screening strategy, LR, SVM, and BPNN models were developed. Receiver operating characteristic (ROC) curves were plotted for each model, and predictive performance was compared using precision, recall, accuracy, F1-score, and area under the ROC curve (AUC).

Results: Significant differences were observed between the two groups in age, BMI, time to medical consultation after onset, length of hospital stay, smoking history, foot ulcer classification, history of toe amputation, ulcer location on the sole, living alone, walking impairment, trauma-induced ulcers, ankle-brachial index, diabetic peripheral neuropathy, lower limb atherosclerosis, multidrug-resistant bacterial infection, glycated hemoglobin, albumin, foot care behavior, and risk perception level ($P<0.05$). Multivariable Logistic regression analysis revealed that length of hospital stay [OR=0.678, 95%CI (0.557, 0.826), $P<0.001$], ulcer location on the sole [OR=0.078, 95%CI (0.011, 0.541), $P=0.010$], living alone [OR=5.689, 95%CI (2.583, 10.726), $P<0.001$], walking impairment [OR=3.364, 95%CI (2.742, 5.638), $P<0.001$], diabetic peripheral neuropathy [OR=3.089, 95%CI (1.156, 4.585), $P=0.025$], lower limb atherosclerosis [OR=4.033, 95%CI (3.688, 9.060), $P<0.001$], multidrug-resistant bacterial infection [OR=3.241, 95%CI (1.728, 7.361), $P<0.001$], glycated hemoglobin

[OR=0.209, 95%CI (0.065, 0.669), P=0.008], albumin [OR=2.038, 95%CI (1.515, 2.741), P<0.001], foot care behavior [OR=1.965, 95%CI (0.874, 3.208), P=0.014], and risk perception level [OR=0.261, 95%CI (0.197, 0.825), P=0.002] were influencing factors for DFU recurrence within one year (P<0.05). In the test set, the accuracy rates for LR, SVM, and BPNN models in predicting DFU recurrence risk were 82.03%, 94.87%, and 87.17%, respectively, with AUCs of 0.843, 0.930, and 0.800. The AUCs of the ROC curves differed significantly among the three models (Z=8.826, P<0.05). The SVM model' s AUC was higher than that of both the LR (Z=5.672, P=0.014) and BPNN models (Z=2.767, P<0.001).

Conclusion: The SVM model demonstrated superior performance in predicting DFU recurrence risk within one year post-discharge, with excellent accuracy, sensitivity, specificity, and AUC. As the relatively optimal model, it warrants further validation and clinical application.

[Keywords] Diabetes mellitus; Foot ulcer; Diabetic foot; Recurrence; Logistic models; Support vector machine; Back propagation neural network; Root cause analysis

Introduction

According to data from the International Diabetes Federation, the global diabetes prevalence was 9.3% in 2019 and is projected to reach 10.9% by 2045 [1]. Concurrently, the prevalence of chronic wounds is increasing globally, with diabetic foot ulcers (DFUs) and lower extremity venous ulcers being the most common types [2]. DFUs represent one of the most severe complications of diabetes, leading to high mortality and disability rates [3]. The lifetime risk for DFU patients reaches 15%-25% [4], with an annual amputation rate of 5.1% [1]. Studies have reported that even after wound healing, DFU recurrence rates reach 40% within one year, 50%-60% within three years, and 65% within five years [4-7]. Clearly, DFU prognosis is unstable, with high recurrence rates and greater early recurrence risk.

Previous studies have established DFU recurrence risk prediction models using various approaches. Lü et al. [8] developed a Logistic regression model identifying smoking, callus formation, abnormal foot skin color, clavus, diabetic peripheral neuropathy, and coronary heart disease as risk factors, with a 26.9% one-year recurrence rate. CRAWFORD et al. [9] found that DFU history, inability to feel a 10g monofilament, and absence of any foot pulses were independent risk factors. More recently, AAN DE STEGGE et al. [10] established two prediction models: one for recurrent plantar ulcers and another for ulcers caused by unrecognized repetitive stress. However, previous research has typically employed single modeling approaches without comparing predictive performance across different algorithms. This study utilizes machine learning algorithms—specifically support vector machine (SVM) and back propagation neural network

(BPNN)—to develop DFU recurrence risk prediction models and compares their predictive efficacy to enable clinical staff to promptly identify at-risk patients and guide preventive treatment decisions.

Methods

1.1 Study Subjects

Patients with DFUs who attended the Department of Burn and Plastic Surgery, Endocrinology Department, and Wound Ostomy Clinic at Ningxia Medical University General Hospital between January 2020 and December 2022 were selected as the study cohort for model development. Inclusion criteria were: (1) diagnosis of type 2 diabetes according to the "Chinese Clinical Guidelines for the Prevention and Treatment of Type 2 Diabetes in the Elderly (2022 Edition)" [11]; (2) initial presentation due to primary DFU; and (3) voluntary participation with signed informed consent. Exclusion criteria included: (1) diabetes with pregnancy or special types of diabetes; (2) severe malignant tumors or other critical illnesses; (3) communication barriers; (4) psychiatric disorders; and (5) loss to follow-up. This study was approved by the Ethics Committee of Ningxia Medical University General Hospital (Approval No.: KYLL-2021-677), and all participants provided informed consent.

1.2 Sample Size Calculation

This study included 36 candidate predictor variables. According to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines for sample size requirements [14], the number of outcome events should be at least 10 times the number of candidate predictors. Therefore, a minimum of 360 DFU outcome events was required. After data cleaning, 390 DFU patients were included. The outcome variable was DFU recurrence within one year post-discharge. DFU recurrence was diagnosed according to the International Working Group on the Diabetic Foot (IWGDF) guidelines [12] and the "2020 IWGDF Guidelines on the Diagnosis of the Diabetic Foot" [13]: a new foot ulcer in a patient with previous DFU history, regardless of location. Patients were followed for one year, starting from the time of initial DFU hospitalization and healing. If a patient experienced multiple recurrences during follow-up, only the first recurrence was recorded, with information collected at the time of first recurrence serving as the reference. Based on DFU recurrence within one year post-discharge, patients were divided into a recurrence group (116 cases, 29.7%) and a non-recurrence group (274 cases, 70.3%).

1.3 Data Collection Instruments

1.3.1 General Data Based on research objectives and content, the research team designed a general data collection form through literature review, preliminary surveys, and expert consultation. The form included three components: (1) Sociodemographic characteristics: gender, age, BMI, time to medical consultation after DFU onset, length of hospital stay, etc.; (2) Medical history assessment: diabetes duration, smoking history ($ $1 cigarette daily for $ 6*monthscumulativelyandcurrentlysmoking*[14]), *alcoholconsumptionhistory*($1 drink weekly for $ $6 months cumulatively and currently drinking [14]), DFU history, hypertension, coronary heart disease, diabetic peripheral neuropathy, diabetic retinopathy, diabetic nephropathy; and (3) Clinical case data: foot ulcer classification, foot skin color, toe involvement, ulcer location on sole, foot callus, history of toe amputation, lower limb atherosclerosis, living alone, walking impairment, trauma-induced ulcer, ankle-brachial index, osteomyelitis, multidrug-resistant bacterial infection, glycated hemoglobin, total bilirubin, albumin, creatinine, cholesterol, hemoglobin, and white blood cell count.

1.3.2 Diabetes Foot Self-care Behavior Scale (DFSBS) The DFSBS, developed by Jin et al. [15], was used to assess diabetes foot self-management behaviors. The scale comprises 7 items. Four items assess the number of days in the past week patients performed self-examination of foot soles, toe spaces, washing between toes, and drying between toes, rated on a 5-point scale (0, 1-2, 3-4, 5-6, and 7 days corresponding to scores 1-5). The remaining three items assess frequency of foot care behaviors ("never," "rarely," "sometimes," "often," "always," scored 1-5). Total scores range from 7 to 35, with higher scores indicating better self-management behaviors. The scale' s Cronbach' s α coefficient is 0.835, with test-retest reliability of 0.916.

1.3.3 Chronic Disease Risk Perception Questionnaire The Chronic Disease Risk Perception Questionnaire, developed by Fang et al. [16], was used to assess patients' DFU risk perception. The questionnaire evaluates three dimensions across 12 items: economic risk, physical diagnosis/treatment risk, and psychosocial risk. Using a 5-point Likert scale ("definitely will not happen," "unlikely," "may happen," "likely," "definitely will happen," scored 1-5), total scores range from 12 to 60, with higher scores indicating greater risk perception. The questionnaire' s overall Cronbach' s α coefficient is 0.833, with subscale coefficients of 0.716 for economic risk, 0.769 for physical diagnosis/treatment risk, and 0.781 for psychosocial risk, demonstrating good reliability and validity.

1.4 Quality Control

Clinical data were collected through the hospital information system. Follow-up data on DFU recurrence were recorded through telephone follow-up, outpatient follow-up, and an off-site follow-up management system, with a follow-up period of one year. Recurrence status and symptom severity were assessed by clinical physicians, diabetes specialist nurses, or wound specialist nurses. Before data collection, three diabetes specialist nurses and three wound specialist nurses (six total) were selected from the Department of Burn and Plastic Surgery, Wound Ostomy Clinic, and Endocrinology Department as data collectors. The researcher provided standardized training on follow-up data collection content, methods, and procedures. A pilot survey of five DFU patients was conducted to refine the data collection form. The "Diabetic Foot Ulcer Questionnaire" was created using Questionnaire Star, and the "Diabetic Foot Ulcer Online Data Collection Form" was developed using Kingsoft Docs for data collection.

1.5 Statistical Analysis

1.5.1 Data Preliminary Screening and Preprocessing Data variable characteristics were reviewed and initially screened according to established rules: (1) variables with >90% missing data per column were deleted; (2) variables with >90% single category proportion per column were deleted; (3) variables with coefficient of variation (CV) <0.05 were deleted; and (4) illogical outliers were removed. Missing values for continuous variables were imputed using median values, while missing values for categorical variables were imputed using mode values.

1.5.2 Feature and Algorithm Selection Feature selection was performed in the training dataset. The outcome variable was DFU recurrence within one year, with other variables as predictors. Initial stepwise Logistic regression analysis was conducted, utilizing OR values and 95%CI to enhance model interpretability. Predictors with P<0.05 in univariate analysis were included in multivariate analysis, with final predictors determined through stepwise regression. BPNN imposes no strict distributional assumptions on data, can identify complex nonlinear relationships between variables, possesses strong adaptive capabilities, and demonstrates high fault tolerance [17]. SVM is a linear and nonlinear classification method with advantages of high prediction reliability, strong stability, and excellent generalization ability [18].

1.5.3 Model Development and Evaluation Eleven significant independent predictors from Logistic regression were included to construct LR, BPNN, and SVM risk prediction models. The training and test sets were split at a 7:3 ratio, with 292 samples for training and 98 for testing. Model evaluation metrics included area under the ROC curve (AUC), precision, accuracy, recall, and F1-score. When metrics were inconsistent, AUC served as the primary reference: AUC of 0.50-0.70 indicated poor predictive performance, 0.70-0.90 indicated moderate performance, and >0.9 indicated excellent performance. SPSS 25.0 and MATLAB R2020b were used for data analysis. Normally distributed quantitative data were expressed as ($\bar{x}\pm s$) and compared between groups using independent samples t-test. Non-normally distributed quantitative data were expressed as M(P25, P75) and compared using Mann-Whitney U test. Categorical data were expressed as frequency and percentage, compared using ^{2} test, with rank data compared using Wilcoxon W test. SPSS 25.0 was used to construct the Logistic regression model, while MATLAB R2020b built the BPNN and SVM models. Model goodness-of-fit was assessed using the Hosmer-Lemeshow test. ROC curves were plotted for each model, and DeLong' s test was used to compare model performance and AUC differences. $P<0.05$ was considered statistically significant.

Results

2.1 Comparison of General Data

No significant differences were found between groups in gender, diabetes duration, alcohol consumption history, foot skin color, foot callus, toe involvement, hypertension history, coronary heart disease history, diabetic retinopathy, diabetic nephropathy history, osteomyelitis, total bilirubin, creatinine, cholesterol, hemoglobin, or white blood cell count ($P>0.05$). However, significant differences were observed in age, BMI, time to medical consultation after onset, length of hospital stay, smoking history, foot ulcer classification, history of toe amputation, ulcer location on the sole, living alone, walking impairment, trauma-induced ulcer, ankle-brachial index, diabetic peripheral neuropathy, lower limb atherosclerosis, multidrug-resistant bacterial infection, glycated hemoglobin, albumin, foot care behavior, and risk perception level ($P<0.05$) .

2.2 Multivariable Logistic Regression Analysis

Using DFU recurrence within one year post-discharge as the dependent variable (no=0, yes=1) and factors showing significant differences in Table 1 as independent variables (see Table 2 for assignments), multivariable Logistic regression analysis revealed that length of hospital stay, ulcer location on the sole, living alone, walking impairment, diabetic peripheral neuropathy, lower limb atherosclerosis, multidrug-resistant bacterial infection, glycated hemoglobin, albumin, foot care behavior, and risk perception level were influencing factors for DFU recurrence ($P<0.05$) .

2.3 Model Performance Results

2.3.1 Logistic Regression Model Results The 11 significant factors from multivariable Logistic regression analysis were used as input variables for the prediction model. In the training set, the model achieved precision, accuracy, and AUC of 83.26%, 81.41%, and 0.825, respectively, when comparing fitted values to actual values.

2.3.2 SVM Model Results Using DFU recurrence as the dependent variable (no=0, yes=1) and variables selected from multivariable Logistic regression analysis as independent variables (assignments same as Table 2), an SVM model was established in the training set. The tune.svm() function with grid search method was used to identify optimal parameters by testing various combinations, with initial C and γ value ranges of (0.01, 0.1, 1, 10, 100). The optimal parameters achieving lowest 10-fold cross-validation error rate were C=100 and γ=0.01. Under these parameters, the SVM model achieved precision, accuracy, and AUC of 97.20%, 96.33%, and 0.940, respectively, in the training set .

2.3.3 BPNN Model Results Using DFU recurrence as the dependent variable (no=0, yes=1) and variables with statistical significance in multivariable Logistic regression analysis as independent variables (assignments same as Table 2), continuous variables were normalized to the (0,1) interval using standardization formulas before input into the BPNN model. The prediction accuracy was 87.17% in the test set and 88.64% in the training set .

2.4 Comparison of the Three Models' Predictive Value

All three models showed high accuracy in predicting DFU recurrence. In both training and test sets, the AUC ranking was: SVM > Logistic regression > BPNN. ROC curves demonstrated that SVM achieved the highest AUC in both training and test sets, with a test set AUC of 0.930 [95%CI (0.916, 0.955)] and highest accuracy of 94.87%. Sensitivity was 0.99 and specificity was 0.87, indicating superior overall performance compared to the other two models. DeLong' s test revealed significant differences in AUCs among the three models ($Z=8.826$, $P<0.05$). The SVM model' s AUC was significantly higher than both Logistic regression ($Z=5.672$, $P=0.014$) and BPNN models ($Z=2.767$, $P<0.001$). The Logistic regression model' s AUC was also significantly higher than BPNN ($Z=3.274$, $P=0.028$) [Figure 1: see original paper], .

Discussion

DFU recurrence is an important marker of poor prognosis in diabetic patients [19]. This study examined subjective and objective variables associated with DFU recurrence, developed an online screening tool for high-risk factors, and used machine learning algorithms to explore risk factors and construct optimal prediction models for short-term follow-up (one year post-discharge). The one-year recurrence rate of 29.74% in this study is higher than the 26.9% reported by Lü et al. [8] but lower than the 31.6% reported by WANG et al. [20] and JIANG et al. [21], and lower than findings from HICKS et al. [22]. Current regional reports of DFU recurrence rates vary [23-24], likely due to differences in medical technology, socioeconomic status, regional diabetes education coverage, high-risk foot screening, and precision treatment. Early screening and prevention for high-risk feet are particularly important for preventing DFU recurrence [25]. Prediction models developed through data mining techniques can help identify DFU recurrence risk factors early in outpatient settings, enabling timely detection of high-risk populations and implementation of interventions. For severe complex DFU cases, rapid referral pathways to specialized foot clinics can be activated, significantly improving population health outcomes.

This study identified influencing factors from both simple personal and complex clinical levels, applied machine learning algorithms to establish optimal prediction models, and conducted comprehensive model evaluation. Multivariable Logistic regression analysis identified length of hospital stay, ulcer location on the sole, living alone, walking impairment, diabetic peripheral neuropathy, lower limb atherosclerosis, multidrug-resistant bacterial infection, glycated hemoglobin, albumin, foot care behavior, and risk perception level as influencing factors for one-year DFU recurrence, consistent with most previous studies [7,8,11,13,24] and providing feasible intervention directions for early screening and prevention.

Many previous DFU recurrence risk studies have used traditional or common machine learning algorithms alone, rarely employing diverse machine learning approaches to compare predictive performance. This study included 36 variables in the initial screening, incorporated 11 statistically significant factors from multivariable Logistic regression into model construction, and developed LR, SVM, and BPNN risk prediction models. Performance was evaluated in the test set, with all three models achieving AUC >0.70 and accuracy >80%, indicating high accuracy and good predictive capability [26-27]. In both training and test sets, SVM outperformed LR and BPNN across all evaluation metrics. BPNN showed lower recall but higher precision compared to the Logistic model. Since single metrics cannot comprehensively evaluate algorithm performance, this study used the F1-score to reconcile precision-recall trade-offs and provide a more accurate comprehensive evaluation [28], with results showing BPNN > Logistic regression, possibly due to sample size and variable count affecting model performance. The limited test set sample size may have restricted BPNN' s algorithmic advantages. Overall, the SVM-based prediction model demonstrated good diagnostic efficacy and stability, with superior accuracy, sensitivity, and AUC compared to the other two models, suggesting that a prediction model combining clinical case data and patient-reported outcomes can serve as an effective auxiliary decision-making tool for predicting DFU recurrence with promising application prospects.

Machine learning algorithms are rapidly advancing in medical diagnosis and health risk monitoring. Prediction models and intelligent health management systems based on machine learning can enable remote monitoring, real-time updates, and intelligent health alerts for disease recurrence screening data [29]. Machine learning algorithms offer significant advantages in acquiring data characteristics and analyzing complex data, evolving from traditional statistical methods like Logistic regression and decision trees to emerging deep neural network algorithms, with widespread applications across fields [30]. SVM is a commonly used machine learning algorithm with powerful learning capabilities. Rao et al. [31] used SVM, BPNN, and Logistic regression to develop breast cancer recurrence prediction models, finding SVM and BPNN showed better predictive performance, similar to our findings. Given the importance of rich clinical DFU data in recurrence risk prediction, effective utilization of these risk prediction models to control high-risk factors in advance could prevent or delay adverse disease progression. Different models have distinct advantages: compared to SVM, BPNN has limitations such as local minima and slow convergence; compared to Logistic regression, SVM does not directly depend on data distribution and offers clear advantages in classification performance, generalization ability, and computational efficiency for multidimensional, nonlinear medical data [32]; while Logistic regression coefficients can explain epidemiological significance of influencing factors, they cannot clearly interpret variables with severe multicollinearity [33-34].

In summary, the SVM-based DFU recurrence risk prediction model demonstrated optimal performance. Its application in clinical DFU assessment will improve screening rates for high-recurrence-risk populations, enable early risk warnings, reduce disease burden, save healthcare costs, and provide clinical guidance for DFU prevention and treatment. Additionally, in clinical research, predicting recurrence risk can facilitate patient stratification to control between-group comparability. As this study only examined short-term prognosis, future research should expand sample sizes, include diverse geographic populations, conduct multicenter large-sample cohort studies, explore new input variables, and optimize model accuracy to further validate predictive performance. Employing multiple machine learning algorithms such as decision trees and random forests for comparison and combined prediction could comprehensively evaluate model predictive value and construct more authoritative algorithms to enhance research accuracy and generalizability.

Author Contributions: ZHANG Juan conceptualized the study, designed the research protocol, conducted feasibility analysis including modeling approaches, collected data, and drafted the manuscript. ZHANG Juan, LI Haifen, LI Xiaoman, YAO Miao, and MA Huizhen were responsible for participant screening, data collection and organization, statistical modeling, results analysis and interpretation, and figure/table creation. MA Qiang revised the manuscript, provided quality control and review, secured funding and material support, and assumed overall responsibility for supervision.

Conflict of Interest: The authors declare no conflicts of interest.

References

[1] SAEEDI P, PETERSOHN I, SALPEA P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the International Diabetes Federation Diabetes Atlas, 9th edition[J]. Diabetes Res Clin Pract, 2019, 157: 107843. DOI: 10.1016/j.diabres.2019.107843.

[2] TENTOLOURIS N, EDMONDS M E, JUDE E B, et al. Editorial: understanding diabetic foot disease: current status and emerging treatment approaches[J]. Front Endocrinol (Lausanne), 2021, 12: 753181. DOI: 10.3389/fendo.2021.753181.

[3] ZHOU X Q, LI M, XIAO M F, et al. ERβ accelerates diabetic wound healing by ameliorating hyperglycemia-induced persistent oxidative stress[J]. Front Endocrinol (Lausanne), 2019, 10: 499. DOI: 10.3389/fendo.2019.00499.

[4] CAVANAGH P R, LIPSKY B A, BRADBURY A W, et al. Treatment for diabetic foot ulcers[J]. Lancet, 2005, 366(9498): 1725-1735. DOI: 10.1016/s0140-6736(05)67699-4.

[5] MCGLOIN H, DEVANE D, MCINTOSH C D, et al. Psychological inter-

ventions for treating foot ulcers, and preventing their recurrence, in people with diabetes[J]. Cochrane Database Syst Rev, 2021, 2(2): CD012835. DOI: 10.1002/14651858.CD012835.pub2.

[6] YADAV C, SRIKANTIAH R M, MANJREKAR P, et al. Assessment of mineral pathophysiology in patients with diabetic foot ulcer[J]. Biol Trace Elem Res, 2020, 195(2): 366-372. DOI: 10.1007/s12011-019-01868-3.

[7] ARMSTRONG D G, BOULTON A J M, BUS S A. Diabetic foot ulcers and their recurrence[J]. N Engl J Med, 2017, 376(24): 2367-2375. DOI: 10.1056/NEJMra1615439.

[8] LU J, YUAN L, LI R, et al. Construction of a recurrence risk prediction model for diabetic foot ulcer[J]. Nursing Research, 2022, 36(6): 993-998. DOI: 10.12102/j.issn.1009-6493.2022.06.009.

[9] CRAWFORD F, CEZARD G, CHAPPELL F M, et al. The development and validation of a multivariable prognostic model to predict foot ulceration in diabetes using a systematic review and individual patient data meta-analyses[J]. Diabet Med, 2018, 35(11): 1480-1493. DOI: 10.1111/dme.13797.

[10] AAN DE STEGGE W B, SCHUT M C, ABU-HANNA A, et al. Development of a prediction model for foot ulcer recurrence in people with diabetes using easy-to-obtain clinical variables[J]. BMJ Open Diabetes Res Care, 2021, 9(1): e002257. DOI: 10.1136/bmjdrc-2021-002257.

[11] Chinese Clinical Guidelines for the Prevention and Treatment of Type 2 Diabetes in the Elderly (2022 Edition) Writing Group. Chinese clinical guidelines for the prevention and treatment of type 2 diabetes in the elderly (2022 edition)[J]. Chinese Journal of Diabetes, 2022, 30(1): 10-11. DOI: 10.3969/j.issn.1006-6187.2022.01.002.

[12] BUS S A, LAVERY L A, MONTEIRO-SOARES M, et al. Guidelines on the prevention of foot ulcers in persons with diabetes (IWGDF 2019 update)[J]. Diabetes Metab Res Rev, 2020, 36(Suppl 1): e3269. DOI: 10.1002/dmrr.3269.

[13] HINCHLIFFE R J, FORSYTHE R O, APELQVIST J, et al. Guidelines of the international writing group on the diabetic foot on diagnosis[J]. Diabetes/Metabolism Research and Reviews, 2020, e3276. DOI: 10.1002/dmrr.3276.

[14] LI B. Construction and validation of a risk prediction model for vitiligo comorbid with diabetes[D]. Xi' an: Air Force Medical University, 2022.

[15] CHIN Y F, HUANG T T. Development and validation of a diabetes foot self-care behavior scale[J]. J Nurs Res, 2013, 21(1): 19-25. DOI: 10.1097/jnr.0b013e3182828e59.

[16] FANG L. Development of the Chronic Disease Risk Perception Questionnaire and study on its influencing factors[D]. Xi' an: Fourth Military Medical University, 2015.

[17] SONG Y, CHEN B, LIU J, et al. Comparison of BP neural network model and Logistic regression model in analyzing influencing factors of recurrence in schizophrenia patients[J]. Journal of Nursing, 2022, 29(13): 59-65. DOI: 10.16460/j.cnki.csa.005181.

[18] MA Q, SUN D, SHI J, et al. Prediction of tumor risk in adult population based on support vector machine and XGboost[J]. Chinese General Practice, 2020(12): 1486-1491. DOI: 10.12114/j.issn.1007-9572.2020.00.066.

[19] PEIKES D, CHEN A, SCHORE J, et al. Effects of care coordination on hospitalization, quality of care, and health care expenditures among Medicare beneficiaries: 15 randomized trials[J]. JAMA, 2009, 301(6): 603-618. DOI: 10.1001/jama.2009.126.

[20] WANG L M, GAO P, ZHANG M, et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013[J]. JAMA, 2017, 317(24): 2515-2523. DOI: 10.1001/jama.2017.7596.

[21] JIANG Y F, WANG X M, XIA L, et al. A cohort study of diabetic patients and diabetic foot ulceration patients in China[J]. Wound Repair Regen, 2015, 23(2): 222-230. DOI: 10.1111/wrr.12263.

[22] HICKS C W, CANNER J K, MATHIOUDAKIS N, et al. Incidence and risk factors associated with ulcer recurrence among patients with diabetic foot ulcers treated in a multidisciplinary setting[J]. J Surg Res, 2020, 246: 243-250. DOI: 10.1016/j.jss.2019.09.025.

[23] FU X L, DING H, MIAO W W, et al. Global recurrence rates in diabetic foot ulcers: a systematic review and meta-analysis[J]. Diabetes Metab Res Rev, 2019, 35(6): e3160. DOI: 10.1002/dmrr.3160.

[24] Professional Committee of Diabetes and Microcirculation of Chinese Society of Microcirculation, Professional Committee of Wound Prevention and Treatment and Tissue Repair of Chinese Research Hospital Association. Expert consensus on primary screening, prevention and treatment of diabetic foot[J]. Chinese Journal of Diabetes, 2019, 27(6): 401-407. DOI: 10.3969/j.issn.1006-6187.2019.06.001.

[25] XIA L, ZHUANG R, WU L, et al. Meta-analysis of risk factors associated with recurrent diabetic foot ulcer[J]. Chinese Journal of Modern Nursing, 2022, 28(9): 1143-1148. DOI: 10.3760/cma.j.cn115682-20210802-03422.

[26] ZHOU L, LIU X, GAO J. Establishment of an artificial neural network prediction model for postoperative fatigue syndrome in patients undergoing painless colonoscopy[J]. Chinese Journal of Anesthesiology, 2021, 41(4): 397-400. DOI: 10.3760/cma.j.cn131073.20201202.00403.

[27] MENG M, GUAN Y, GUO L, et al. Construction and validation of a risk prediction model for post-intensive care syndrome in patients after cardiovascular surgery[J]. Chinese Journal of Nursing, 2022, 57(12): 1486-1494. DOI: 10.3761/j.issn.0254-1769.2022.12.006.

[28] WANG J, YANG Y, ZHENG Z, et al. Comparison of three risk prediction models for metabolic syndrome in oil workers[J]. Modern Preventive Medicine, 2021, 48(4): 613-618.

[29] ZHOU L, WEN X, JIANG R, et al. Research progress on application of machine learning in nursing[J]. Journal of Nurses Training, 2022, 37(15): 1388-1392. DOI: 10.16821/j.cnki.hsjx.2022.15.009.

[30] WU Q, LU Z, LIU Y, et al. Visual analysis of application progress of machine learning methods in critical care medicine from 2010 to 2020[J]. Chinese Journal of Critical Care Medicine, 2021, 41(11): 959-965. DOI: 10.3969/j.issn.1002-1949.2021.11.001.

[31] RAO P, YE F. Study on influencing factors of breast cancer recurrence based on Logistic regression, ANN and SVM[J]. Computer Systems & Applications, 2016, 25(7): 259-263. DOI: 10.15888/j.cnki.csa.005181.

[32] ZHAN J B, WANG Q, CHENG J, et al. Current status of severe fever with thrombocytopenia syndrome in China[J]. Virol Sin, 2017, 32(1): 51-62. DOI: 10.1007/s12250-016-3931-1.

[33] WANG Q, LIU S. Construction of a multi-layer artificial neural network classification prediction model for subclinical atherosclerosis in patients with type 2 diabetes[J]. Chinese General Practice, 2021, 24(36): 4612-4617. DOI: 10.12114/j.issn.1007-9572.2021.00.066.

[34] QIN W, GAO M, SHEN Y, et al. 3-month blood glucose prediction in patients with type 2 diabetes based on machine learning algorithms[J]. Chinese Journal of Disease Control & Prevention, 2019, 23(11): 1313-1317. DOI: 10.16462/j.cnki.zhjbkz.2019.11.003.

(Received: March 22, 2023; Revised: May 23, 2023)

(Editor: KANG Yanhui)

Note: Figure translations are in progress. See original paper for figures.