Am J Transl Res 2026;18(2):1690-1704
www.ajtr.org /ISSN:1943-8141/AJTR0169981

Original Article

Construction of an auxiliary diagnostic model for secondary pulmonary bacterial infection in influenza based on routine detection indicators

Li-Zhi He, Guan-Cheng Huang, Yun-Feng Bao

Department of Infectious Diseases, The Affiliated Yangming Hospital of Ningbo University (Yuyao People's Hospital), Yuyao 315400, Zhejiang, China

Received November 13, 2025; Accepted January 31, 2026; Epub February 15, 2026; Published February 28, 2026

Abstract: Objective: To explore the feasibility of establishing an auxiliary diagnostic model for secondary pulmonary bacterial infection in influenza using routine indicators from fever clinics. Methods: A retrospective analysis of 510 influenza cases (divided into modeling and validation sets in a 7:3 ratio) was conducted, with an additional 72 cases selected for external validation. Logistic regression was adopted as the traditional diagnostic model, while two machine learning models (decision tree and random forest) were constructed using R4.2.3 software. The diagnostic performance of each model was evaluated using multiple indicators, including accuracy, sensitivity, specificity, positive predictive value, negative predictive value, receiver operating characteristic curve, and area under the curve (AUC). Results: Among the 357 influenza patients in the modeling set, 101 developed secondary pulmonary bacterial infection, while 256 did not. Multivariate logistic regression analysis showed that age, white blood cell count, C-reactive protein, serum amyloid A, creatine kinase isoenzyme, and D-dimer were independent risk factors for secondary pulmonary bacterial infection (all $P<0.05$). In the modeling set, validation set, and external validation set, the machine learning model generally outperformed the logistic regression model in all diagnostic performance metrics. The random forest model performed exceptionally well on all three datasets, with AUC values of 0.951, 0.902, and 0.852, respectively. Conclusion: In auxiliary diagnostic models constructed based on routine fever clinic testing indicators, machine learning models, especially the random forest model, demonstrate high diagnostic accuracy and good generalization ability for influenza-related secondary pulmonary bacterial infection.

Keywords: Influenza, pulmonary bacterial infection, auxiliary diagnostic model, routine detection indicators, machine learning

Introduction

Influenza (flu) is an acute respiratory infection caused by the influenza virus [1]. It has a strong ability to spread and infect. After infection, patients' immunity usually decreases, which increases their susceptibility to secondary bacterial pneumonia [2]. This complication can accelerate disease progression, complicate treatment, and increase patient mortality [3]. According to the results of a recent large-scale epidemiological study, patients with influenza complicated by bacterial pneumonia have longer hospital stays and a mortality rate three times higher than those with influenza alone [4]. Timely and accurate diagnosis of secondary bacterial pneumonia caused by influenza is crucial for timely implementation of optimal treatment measures and improvement of recovery. Routine examinations are performed in fever clinics, which provide a wealth of data for clinical diagnosis. Nevertheless, it remains difficult to build auxiliary diagnostic models using these traditional indicators. Although previous studies have investigated single indicators or simple combinations of indicators, their diagnostic accuracy and efficiency still do not meet clinical needs [5]. This study innovatively combines a series of routine test indicators from fever clinics. By collecting case data and directly comparing traditional diagnostic models with machine learning techniques, we hope to build a robust auxiliary diagnostic model. Machine learning models, including decision trees and

https://doi.org/10.62347/KLBR4323

random forests, can automatically identify complex patterns and correlations in data. This capability may help overcome the shortcomings of traditional diagnostic tools while improving the accuracy and effectiveness of diagnosis. This study not only provides new insights into the early diagnosis of influenza-associated bacterial pneumonia but also shows great promise for clinical application, which will contribute to improving the overall management of influenza cases.

Materials and methods

Study subjects

We studied 510 influenza patients at Affiliated Yangming Hospital of Ningbo University (Yuyao People's Hospital) over two periods: November 2023 to April 2024 and December 2024 to March 2025. The sample size was calculated based on the diagnostic model. Studies have indicated that the ratio of total sample size to the number of predictive variables should be at least 10:1 [6, 7]. This study initially included 28 routine diagnostic indicators as candidate predictive variables, thus requiring a minimum sample size of 280 cases. Considering a potential 20% data loss rate due to incomplete follow-up or missing data, the target sample size was expanded to 336 cases. The final sample size of 510 cases exceeded the computational requirements, ensuring sufficient statistical power for model construction and validation.

These cases were divided into a modeling set (357 cases) and a validation set (153 cases) at a 7:3 ratio.

Inclusion criteria were as follows: (1) Confirmed influenza virus infection by reverse transcription-polymerase chain reaction or immunochromatography using nasal or pharyngeal swab specimens [8]; (2) Aged >18 years; (3) Complete routine laboratory and imaging examination results from the fever clinic; (4) Clear clinical diagnosis of secondary pulmonary bacterial infection [9, 10].

Exclusion criteria included: (1) Comorbid severe underlying diseases that may affect the test results or diagnosis (such as advanced malignant tumors, end-stage autoimmune diseases) [during the study period, the proportion of influenza patients with serious underlying comorbidities among those who visited the fever clinic was 15.28% (92 out of 602 patients)]; (2) Missing or incomplete test data.

In addition, 72 influenza patients who met the inclusion criteria during the same period (data collection time from April to September 2025) were selected as an external validation set to assess the generalization ability of the model under different data settings [11, 12]. All 72 patients were from the Affiliated Yangming Hospital of Ningbo University (Yuyao People's Hospital).

This study was approved by the Medical Ethics Committee of the Affiliated Yangming Hospital of Ningbo University (Yuyao People's Hospital) (approval No.: 2025-07-004).

Data collection

We extracted relevant test data and clinical diagnostic information from cases meeting the inclusion criteria through the hospital information system, covering the following three categories: (1) Baseline data: Age, gender, chronic obstructive pulmonary disease (COPD), coronary heart disease, hypertension, diabetes mellitus, fever, cough, expectoration, body temperature, heart rate, respiratory rate, systolic blood pressure, diastolic blood pressure, oxygen saturation, and hemoglobin. (2) Laboratory tests: Complete blood count - white blood cell count (WBC), C-reactive protein (CRP), serum amyloid A (SAA), procalcitonin (PCT); emergency liver function [alanine transaminase (ALT), aspartate transaminase (AST), total bilirubin (TBIL)]; creatine kinase (CK); emergency routine biochemistry [serum creatinine (Scr), blood urea nitrogen (BUN), potassium (K), sodium (Na)]; myocardial injury markers [troponin I (cTnI), creatine kinase isoenzyme (CK-MB)]; routine coagulation function [prothrombin time (PT), activated partial thromboplastin time (APTT)]; D-dimer; respiratory pathogen nucleic acid detection; influenza A and B virus antigen detection; mycoplasma pneumoniae antibody (MP-Ig) titer; mycoplasma pneumoniae IgM antibody (MP-IgM), mycoplasma pneumoniae RNA (MP RNA); novel coronavirus antigen detection; novel coronavirus RNA detection. (3) Imaging examination: Plain chest CT scan.

Definition and detection of relevant indicators

Blood cell analysis: WBC count: All blood samples were collected within 24 hours of patient

admission, diluted and pre-treated with hemolysis, and then counted individually using flow cytometry based on cell characteristics as they passed through the detection area. This test was performed using a Sysmex XN-9000 fully automated blood analyzer.

Detection of inflammatory markers: CRP and SAA: Both indicators were tested within 24 hours of patient admission using a PreciControl PA 1500 automatic biochemical analyzer. The detection method was immunoturbidimetric, measuring the turbidity of the immune complexes formed by the antigen-antibody reaction.

PCT: PCT was assayed within 24 hours of patient admission using an ETHealth fluorescence analyzer with immunofluorescence. This method quantifies PCT in the sample by the binding of a specific antibody to the procalcitonin antigen.

Pathogen detection: Nucleic acid detection (respiratory pathogen nucleic acid, novel coronavirus RNA, etc.): This detection was performed within 24 hours of patient admission using real-time quantitative PCR. Primers and probes were designed, and nucleic acid amplification was monitored via fluorescence signals. The instrument used in this study was an ABI 7500 real-time quantitative PCR instrument.

Antigen and antibody detection (influenza A/B antigens, mycoplasma pneumoniae IgM antibodies, etc.): Influenza A/B antigens were tested using immunochromatography within 24 hours of patient admission. Mycoplasma pneumoniae IgM antibodies were assayed in serum samples using immunofluorescence within 48 hours of patient admission.

Other blood biochemical and related tests: Biochemical indicators (ALT, Scr, etc.): These indicators were detected within 24 hours of patient admission using a VITROS 5600 automatic biochemical immunoassay analyzer. Detection employed various methods, including rate methods, colorimetric methods, and enzymatic methods, and was achieved by measuring the absorbance produced by the chemical reaction. Myocardial related indicators (CK, cTnI, etc.): These were measured using an immunochemiluminescence assay within 24 hours of patient admission, employing a Beckman Coulter DxI 800 chemiluminescence immunoassay analyzer.

Coagulation indicators (PT, etc.): Plasma coagulation time was measured using a Stago STA-R Max automatic coagulation analyzer (magnetic bead coagulation method) within 24 hours of patient admission. D-dimer was also measured on this analyzer using an immunoturbidimetric method.

Grouping

We divided patients into two groups according to whether they had secondary pulmonary bacterial infection:

(1) Pulmonary bacterial infection group: Confirmation of bacterial pathogens - the core and mandatory diagnostic criterion for bacterial infection, requiring at least one of the following microbiological test results: ① Positive bacterial culture results from lower respiratory tract samples (e.g., bronchoalveolar lavage fluid, sputum), and isolation of pathogenic bacteria (such as Streptococcus pneumoniae, Staphylococcus aureus, Klebsiella pneumoniae); ② Detection of bacterial nucleic acid in lower respiratory tract specimens via nucleic acid amplification tests (NAATs); ③ Positive blood bacterial culture (indicating bacteremia caused by pulmonary bacterial infection). To prevent misclassification, we excluded patients with positive nucleic acid tests for other respiratory viruses (e.g., respiratory syncytial virus, adenovirus) in lower respiratory tract samples (to rule out viral pneumonia), and patients diagnosed with non-infectious inflammatory diseases (such as interstitial lung disease, pulmonary edema, pulmonary embolism) based on clinical data, laboratory tests (such as D-dimer, cardiac function indicators), and imaging characteristics (these diseases may present similarly to bacterial pneumonia). In addition to pathogen confirmation, patients must also possess the following clinical and imaging evidence: typical chest imaging (chest CT scan) findings, including newly appearing infiltrates, consolidations, cavitation, or pleural effusion consistent with bacterial pneumonia; clinical manifestations suggestive of bacterial pneumonia, such as persistent or worsening fever, worsening cough with purulent sputum, dyspnea, and systemic inflammatory signs (e.g., significantly elevated levels of inflammatory markers such as C-reactive protein, serum amyloid A, and procalcitonin).

(2) Influenza group: Patients diagnosed with influenza (confirmed by nasopharyngeal swab reverse transcription-polymerase chain reaction (RT-PCR) or positive for influenza A/B antigen) and not meeting the above criteria for secondary pulmonary bacterial infection were included in this group. These patients may present with typical influenza-related symptoms such as fever, cough, sore throat, myalgia, and fatigue, but must meet strict microbiological and exclusion criteria: ① no confirmed bacterial pathogens (negative respiratory/blood bacterial culture, negative lower respiratory tract bacterial nucleic acid test); ② exclusion of other viral pneumonias (negative nucleic acid test for other common respiratory viruses); ③ exclusion of non-infectious inflammatory lung diseases (via comprehensive assessment of clinical data, relevant laboratory tests, and imaging features); ④ no evidence of other concurrent bacterial lung infections.

Statistical methods

Statistical analysis was performed using SPSS 26.0 software. Non-normally distributed continuous variables were described using medians (interquartile range) [P50 (P25, P75)], and the Mann-Whitney U nonparametric test was used for intergroup comparisons. Normally distributed continuous variables were expressed as mean ± standard deviation (x ± sd), and independent samples t-tests were used for intergroup differences. Categorical variables were presented as counts and percentages [n (%)], and the chi-square test was used for analysis; when the expected frequency was less than 5, the exact probability method was used instead. In addition, we systematically explored the potential influencing factors of secondary pulmonary bacterial infection in influenza patients using a multivariate logistic regression model.

In terms of machine learning model construction, the independent influencing factors identified by multivariate logistic regression were used as predictor variables. Decision tree and random forest models were constructed using R 4.2.3 software, with the decision tree model constructed using the rpart package and the random forest model constructed using the randomForest package. Parameters were optimized using 5-fold stratified cross-validation, stratified by infection/non-infection ratio to ensure the representativeness of each subset. Key parameters were identified and optimized during the study. Cost complexity pruning was performed on the decision tree. The optimal complexity parameter was determined by minimizing the cross-validation error to avoid overfitting. The maximum tree depth was set to 10. The minimum number of leaf node samples was limited to 5. When determining the optimal parameters for the random forest, ntree was optimized to 500 to avoid excessive computation while ensuring convergence. Mtry was set to the square root of the total number of predictors to balance the diversity of the base learners and the predictive accuracy.

We evaluated the model performance using multiple metrics, including accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), receiver operating characteristic (ROC) curve, and area under the curve (AUC). Higher values of these metrics indicate better predictive performance. We also plotted ROC curves to quantify the model's efficacy in predicting the risk of influenza-related secondary pulmonary bacterial infection and used the Z-test to compare the differences in AUC between different models. In all analyses, a two-sided p-value <0.05 was considered statistically significant.

Results

Univariate analysis of influenza-induced secondary pulmonary bacterial infections

Of the 357 patients in the modeling set, 176 were male and 181 were female, aged 19 to 83 years, with a mean age of 45.45±12.57 years. Among them, 101 cases were influenza complicated with secondary pulmonary bacterial infection, and 256 cases were influenza without this complication. Compared with the influenza group, the pulmonary bacterial infection group had higher levels of age, WBC, CRP, SAA, PCT, CK-MB, and D-dimer, and other indicators. The positive rates of respiratory pathogen nucleic acid detection and MP-IgM were also higher, as was the incidence of pulmonary infiltration (all P<0.05) (**Tables 1** and **2**).

Multivariate logistic regression analysis of influenza-induced secondary pulmonary bacterial infections

We constructed a multivariate logistic regression model, using the presence of secondary

Table 1. Comparison of baseline characteristics between the pulmonary bacterial infection group and the influenza group

Baseline Characteristics	Pulmonary bacterial infection group (n = 101)	Influenza group (n = 256)	t/Z/χ^2	P
Age (x±sd, years)	51.38±10.03	45.57±12.29	4.227	<0.001
Gender (Male/Female, cases)	56/45	120/136	2.128	0.145
COPD (Yes/No, cases)	7/94	14/242	0.280	0.597
Coronary heart disease (Yes/No, cases)	9/92	19/237	0.222	0.637
Hypertension (Yes/No, cases)	29/72	65/191	0.412	0.521
Diabetes mellitus (Yes/No, cases)	9/92	18/238	0.366	0.545
Fever (Yes/No, cases)	74/27	185/71	0.036	0.849
Cough (Yes/No, cases)	83/18	213/43	0.054	0.817
Sputum production (Yes/No, cases)	73/28	180/76	0.135	0.713
Body temperature (x±sd, °C)	37.47±0.81	37.47±0.84	<0.001	>0.999
Heart rate (x±sd, beats/min)	102.51±16.67	105.50±16.27	1.553	0.121
Respiratory rate (x±sd, breaths/min)	25.72±4.19	26.51±4.11	1.627	0.105
Systolic blood pressure (x±s, mmHg)	118.59±18.00	119.42±16.49	0.417	0.677
Diastolic blood pressure (x±s, mmHg)	76.63±11.80	74.84±11.38	1.325	0.186
Oxygen saturation (x±sd, %)	94.95±2.60	94.53±2.87	1.278	0.202
Hemoglobin (x±sd, g/L)	129.49±19.99	133.69±17.69	1.946	0.052

COPD: Chronic obstructive pulmonary disease.

pulmonary bacterial infection in influenza patients as the dependent variable (yes = 1; no = 0), and indicators showing significant differences in univariate analysis as independent variables. These variables were as follows: age (actual value), WBC (actual value), CRP (actual value), SAA (actual value), PCT (actual value), respiratory pathogen nucleic acid detection (positive = 1; negative = 0), MP-IgM (positive = 1; negative = 0), pulmonary infiltration (positive = 1; negative = 0), CK-MB, and D-dimer (actual value).

The results showed that age, WBC, CRP, SAA, CK-MB, and D-dimer were independent risk factors for secondary pulmonary bacterial infection in influenza patients (all P<0.05) (**Table 3**).

Model construction

(1) Logistic regression model: We constructed a predictive model for secondary pulmonary bacterial infection induced by influenza using the regression coefficients and constant term obtained from logistic regression analysis. The model formula was as follows:

P=1/e^-Y

Where Y = -16.605 + 0.053 × Age (years) + 0.509 × WBC (×10^9/L) + 0.059 × CRP (mg/L) + 0.052 × SAA (mg/L) + 0.088 × CK-MB (U/L) + 3.097 × D-dimer (mg/L).

Here, P denotes the probability of secondary pulmonary bacterial infection in influenza, and patients' age, WBC, CRP, SAA, CK-MB, and D-dimer are all actual measured values.

We calculated the predicted probability of secondary pulmonary bacterial infection for each influenza patient via this model. The AUC of the model was 0.878, significantly higher than the AUC of any single indicator detected individually. The standard error was 0.021 (P<0.001), and the 95% confidence interval (CI) was 0.866-0.950 (**Table 4** and **Figure 1**). These results confirm that age, WBC, CRP, SAA, CK-MB, and D-dimer are valuable indicators for predicting secondary pulmonary bacterial infection induced by influenza.

(2) Machine learning models: Using independent risk factors (age, WBC, CRP, SAA, CK-MB, D-dimer) as predictor variables, we constructed two machine learning models (decision tree and random forest) using data from the modeling set (**Figures 2** and **3**). (1) For the decision tree model, the core splitting nodes were ranked in the order of SAA (threshold = 41), D-dimer (threshold = 0.48) and CRP (threshold

Table 2. Comparison of laboratory indicators between the pulmonary bacterial infection group and the influenza group

Laboratory Parameters	Pulmonary bacterial infection group (n = 101)	Influenza group (n = 256)	t/Z/χ^2	*P*
The WBC in the blood routine test (×10^9/L)	8.43±1.91	7.11±1.57	6.716	<0.001
CRP (mg/L)	47.77±22.47	32.42±7.74	9.598	<0.001
SAA (mg/L)	45.34 (32.33, 74.01)	29.45±10.90	-7.425	<0.001
PCT (ng/mL)	0.65 (0.40, 0.90)	0.44±0.08	-6.231	<0.001
Positive rate of respiratory pathogen nucleic acid detection (%)	69 (68.32%)	136 (53.13%)	6.837	0.009
Positive rate of MP-IgM (%)	30 (29.70%)	50 (19.53%)	4.310	0.038
ALT (U/L)	31.09±10.32	29.52±8.89	1.434	0.152
AST (U/L)	31.50±12.37	30.30±10.26	0.937	0.349
TBIL (μmol/L)	15.22±5.03	14.63±4.05	1.155	0.249
CK (U/L)	80.07±28.81	79.93±25.59	0.045	0.964
Scr (μmol/L)	79.09±15.77	78.55±9.90	0.388	0.698
Blood BUN (mmol/L)	5.05±1.09	4.83±1.14	1.663	0.097
K (mmol/L)	4.06±0.32	3.96±0.48	1.930	0.054
Na (mmol/L)	140.36±3.08	139.58±4.08	1.736	0.084
cTnI (ng/mL)	0.06±0.01	0.06±0.02	<0.001	>0.999
CK-MB (U/L)	24.98±9.03	18.32±12.27	4.950	<0.001
PT (s)	12.35±1.07	12.23±0.83	1.130	0.259
APTT (s)	29.98±3.02	29.50±2.77	1.437	0.152
D-dimer (mg/L)	0.52 (0.34, 0.64)	0.29±0.09	-9.284	<0.001
Positive rate of influenza A/B antigen detection (%)	50 (49.50%)	105 (41.02%)	2.125	0.145
Positive rate of MP-Ig titre (%)	28 (27.72%)	62 (24.22%)	0.472	0.492
Positive rate of MP RNA (%)	22 (21.78%)	50 (19.53%)	0.228	0.633
Positive rate of SARS-CoV-2 antigen detection (%)	5 (4.95%)	10 (3.91%)	0.196	0.658
Positive rate of SARS-CoV-2 RNA detection (%)	3 (2.97%)	5 (1.95%)	0.035	0.851
Positive rate of pulmonary infiltration (%)	101 (100.00%)	226 (88.28%)	21.020	<0.001

WBC: white blood cells; CRP: C-reactive protein; SAA: serum amyloid A; PCT: procalcitonin; MP-IgM: mycoplasma pneumoniae IgM antibody; ALT: alanine transaminase, AST: aspartate transaminase; TBIL: total bilirubin; CK: creatine kinase; Scr: serum creatinine; BUN: urea nitrogen; K: potassium; Na: sodium; cTnI: troponin I; CK-MB: creatine kinase-myocardial band; PT: prothrombin time; APTT: activated partial thromboplastin time; MP-Ig: mycoplasma pneumoniae antibody; MP: mycoplasma pneumoniae; RNA: ribose nucleic acid; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Table 3. Multivariate Logistic regression analysis of influenza-induced secondary pulmonary bacterial infections

Variable	B	SE	Wals	P	OR (95% CI)
Age	0.053	0.016	11.709	0.001	1.055 (1.023, 1.087)
WBC	0.509	0.110	21.317	<0.001	1.664 (1.340, 2.065)
CRP	0.059	0.014	17.428	<0.001	1.060 (1.032, 1.090)
SAA	0.052	0.013	15.489	<0.001	1.054 (1.027, 1.082)
PCT	2.185	1.304	2.809	0.094	8.891 (0.690, 114.498)
Positive rate of respiratory pathogen nucleic acid detection	0.108	0.355	0.093	0.760	1.114 (0.556, 2.232)
Positive rate of MP-IgM	0.603	0.395	2.327	0.127	1.828 (0.842, 3.968)
Positive rate of pulmonary infiltration	0.641	0.387	2.737	0.098	1.898 (0.888, 4.055)
CK-MB	0.088	0.023	14.499	<0.001	1.092 (1.044, 1.143)
D-dimer	3.097	1.409	4.830	0.028	22.122 (1.398, 350.087)
Constant	-16.605	1.944	72.940	<0.001	0.000

WBC: white blood cells; CRP: C-reactive protein; SAA: serum amyloid protein A; PCT: procalcitonin; MP-IgM: mycoplasma pneumoniae antibody IgM; IgM: immunoglobulin m; CK-MB: creatine kinase-myocardial band; SE: standard error; OR: odds ratio; CI: confidence interval.

= 48). Specifically, when SAA level was ≥41, the model directly classified the case as positive with a predicted probability of 1.00. When the SAA level was <41, it was further stratified according to the D-dimer threshold of 0.48, and the positive predicted probability when D-dimer ≥0.48 was 0.89. For samples with D-dimer <0.48, they were further

Table 4. Logistic regression model

Variable	AUC (95% CI)	SE	*P*
Age	0.643 (0.582, 0.703)	0.031	<0.001
WBC	0.696 (0.634, 0.758)	0.032	<0.001
CRP	0.701 (0.625, 0.776)	0.039	<0.001
SAA	0.711 (0.639, 0.782)	0.036	<0.001
CK-MB	0.669 (0.600, 0.738)	0.035	<0.001
D-dimer	0.783 (0.723, 0.843)	0.031	<0.001
Predict probability	0.878 (0.866, 0.950)	0.021	<0.001

WBC: white blood cells; CRP: C-reactive protein; SAA: serum amyloid protein A; CK-MB: creatine kinase-myocardial band; AUC: area under the curve; CI: confidence interval; SE: standard error.

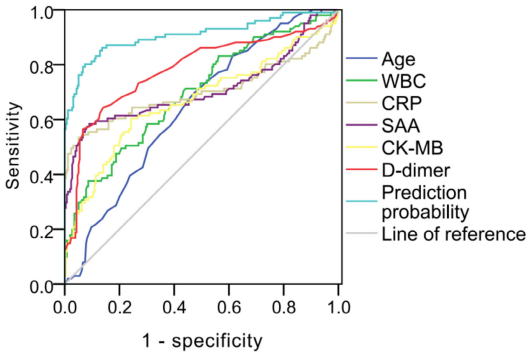

Figure 1. Receiver operator characteristic curve (ROC) of the Logistic regression model. WBC: white blood cells; CRP: C-reactive protein; SAA: serum amyloid protein A; CK-MB: creatine kinase-myocardial band.

classified according to the CRP threshold of 48, and the positive predicted probability when CRP ≥48 was 0.57, and when CRP <48 was 0.07. The sample proportion corresponding to each node (e.g., samples with SAA <41 accounted for 84% of the modeling set samples) also reflected the actual data distribution underlying feature splitting. (2) Feature importance analysis of the random forest model indicated that SAA contributed the most to the model's predictive accuracy (indicated by the largest decrease in average accuracy). The importance of D-dimer and CRP decreased sequentially. In terms of Gini coefficient gain, SAA also showed the largest decrease in average Gini coefficient. These findings suggest that SAA serves as the key feature with the strongest predictive power in this study, consistent with the core splitting node features identified in the decision tree model (**Table 5**).

The ROC curves predicted by the models are shown in **Figure 4**. We tested the predictive performance of both models using validation set data, and the corresponding ROC curves are shown in **Figure 5**.

Comparison of diagnostic efficacy among models

The diagnostic efficacy of the three models on the modeling set, validation set, and external validation set is summarized in **Table 6**. Overall, the two machine learning models (decision tree and random forest) outperformed the logistic regression model on most evaluation metrics, with the random forest model performing better in most cases.

We used the Z-test to compare the AUC values of the three models, with the following results: Compared with the logistic regression model:

Decision tree model: AUC was significantly higher in the modeling set (Z = 2.78, P = 0.005) and validation set (Z = 2.36, P = 0.018), but no significant difference was observed in the external validation set (Z = 1.71, P = 0.087).

Random forest model: AUC was significantly higher in all three sets (modeling set: Z = 5.02<0.001; validation set: Z = 4.37<0.001; external validation set: Z = 3.19, P = 0.001).

When compared with the decision tree model: Random forest model: The AUC was significantly higher on the modeling set (Z = 2.24, P = 0.025) and the validation set (Z = 2.01, P = 0.044), but no significant difference was found on the external validation set (Z = 1.48, P = 0.139).

Optimal cut-off values and corresponding sensitivity/specificity of each model: To determine the optimal cut-off values for each model, we adopted the Youden index maximization principle. The optimal cut-off values and corresponding performance metrics are as follows:

Logistic regression model: Modeling set: Cut-off = 0.42, sensitivity = 0.755, specificity = 0.802; Validation set: Cut-off = 0.45, sensitivity = 0.732, specificity = 0.784; External vali-

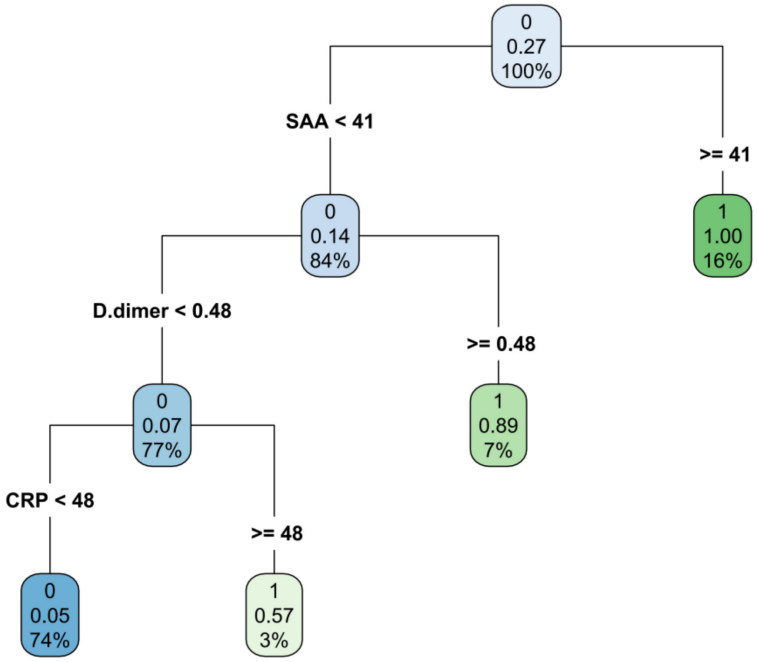

Figure 2. Decision tree model. CRP: C-reactive protein; SAA: serum amyloid protein A.

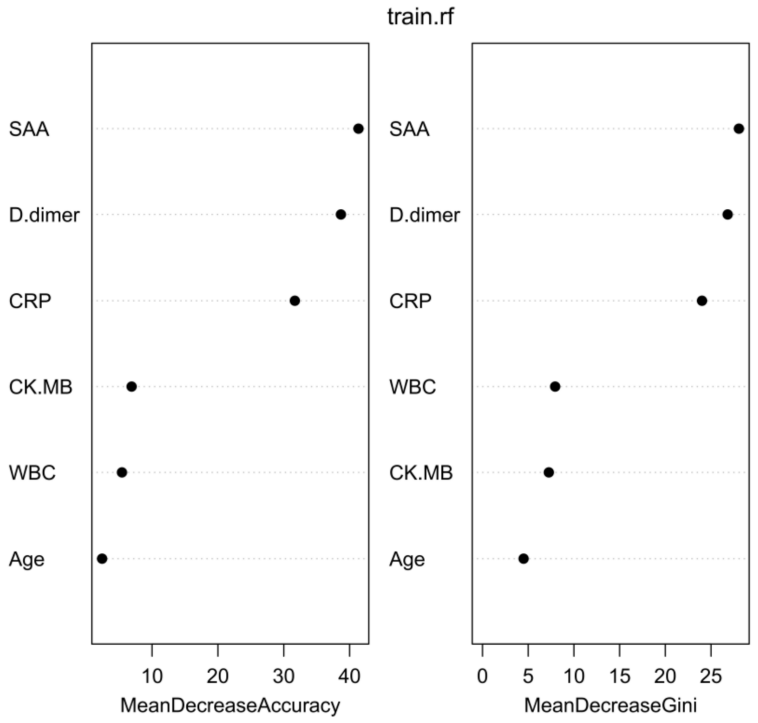

Figure 3. Random forest model. WBC: white blood cells; CRP: C-reactive protein; SAA: serum amyloid protein A; CK-MB: creatine kinase-myocardial band.

dation set: Cut-off = 0.41, sensitivity = 0.711, specificity = 0.765.

Decision tree model: Modeling set: Cut-off = 0.48, sensitivity = 0.868, specificity = 0.972; Validation set: Cut-off = 0.46, sensitivity = 0.758, specificity = 0.987; External validation set: Cut-off = 0.43, sensitivity = 0.771, specificity = 0.803.

Random forest model: Modeling set: Cut-off = 0.50, sensitivity = 0.824, specificity = 0.983; Validation set: Cut-off = 0.49, sensitivity = 0.909, specificity = 0.880; External validation set: Cut-off = 0.47, sensitivity = 0.802, specificity = 0.832.

Model calibration analysis

To evaluate the consistency between the predicted probabilities of the models and the actual infection probabilities, a calibration analysis was performed on each model (**Figure 6**), and the Hosmer-Lemeshow test was used to quantify the calibration performance:

Logistic regression model: In the training set, the Hosmer-Lemeshow test yielded χ^2 = 10.26 and P = 0.179; In the validation set, χ^2 = 9.12 and P = 0.082. These results indicate that the model's predicted probabilities were in good agreement with the actual probabilities.

Decision tree model: In the training set, the Hosmer-Lemeshow test yielded χ^2 = 10.35, P = 0.241; In the validation set, χ^2 = 11.07, P = 0.199. This suggests that the calibration performance of this model is acceptable.

Random forest model: The Hosmer-Lemeshow test results were as follows: In training set, χ^2 = 6.83, P = 0.554; In validation set, χ^2 = 7.51, P = 0.482. This model showed the best calibration performance am-

Table 5. Feature importance results of random forest model

Variable	Mean Decrease Accuracy	Mean Decrease Gini
SAA	39.1912	27.727649
CK-MB	11.722451	10.966733
CRP	20.139722	17.421539
WBC	9.431013	9.111622
D-dimer	29.939817	22.724178
Age	3.927892	5.380894

SAA: serum amyloid protein A; CK-MB: creatine kinase-myocardial band; CRP: C-reactive protein; WBC: white blood cells.

ong the three models, and its predicted probabilities thus had more stable accuracy.

Discussion

Research findings

We successfully developed an auxiliary diagnostic model for influenza-associated secondary bacterial pulmonary infections based on routine clinical testing indicators. Further multivariate logistic regression analysis revealed that age, WBC, CRP, SAA, CK-MB and D-dimer were independent risk factors for these secondary infections. The logistic regression model built based on these risk factors achieved an AUC of 0.878. This model significantly outperformed single-marker detection, confirming its practicality in predicting influenza-related secondary bacterial pulmonary infections. We used decision trees and random forests from machine learning methods to extend the analysis. Subsequent results showed that both machine learning models outperformed the logistic regression model on most diagnostic efficacy metrics. The random forest prediction model developed in this study demonstrated the best overall diagnostic performance across the modeling, validation, and external validation datasets, with AUC values of 0.951, 0.902, and 0.852, respectively. The model's advantages were most evident on the modeling and validation sets. The decision tree models demonstrated excellent specificity on both the modeling and validation sets, but suffered from poor stability on these datasets. Furthermore, their performance on the external validation set is limited due to their significantly weak stability. Nevertheless, the model had a robust specificity during both the modeling and validation phases. Logistic regression models performed the worst across all metrics; their accuracy and AUC droped even more significantly when extrapolating the dataset, indicating poor generalization ability.

Comparison with recent studies

Compared with previous studies, our research has distinctive advantages in indicator selection and model construction. In terms of indicator selection, most existing studies only focus on single or a few inflammatory markers, and there is a gap in the comprehensive application of multidimensional routine clinical indicators to improve predictive accuracy. For example, Zhou et al. [13] reported that PCT is an effective indicator for predicting secondary bacterial infection in elderly COVID-19 patients, but they ignored the synergistic effect of other routine clinical indicators. Campani et al. [14] evaluated the value of PCT and CRP in identifying secondary infection in critically ill COVID-19 patients in the ICU, and found that single or a few inflammatory markers have limitations and more routine clinical indicators were not included. Yang et al. [15] explored the role of markers such as PCT and interleukin-6 in the diagnosis of bloodstream infection, but failed to further explore the comprehensive synergistic predictive value of multiple indicators. A quantitative comparison with the core performance metrics in recent literature highlights the advantages of our model: a study on the prediction of severe influenza A-related illness in children conducted by Xiong et al. [16] used a logistic regression model containing five clinical and laboratory variables (such as the neutrophil-to-lymphocyte ratio and hemoglobin), and reported an AUC of 0.905 in the training set, which is comparable to the AUC of 0.852 for the external validation of our random forest model. It is worth noting that their study focused on predicting severe influenza A infection in children, while our study focused on influenza-associated secondary bacterial pulmonary infection in adults with a wider age distribution, which also explained the subtle difference in performance between the two. Another multicenter study conducted by Hu et al. [17] developed a CatBoost model for influenza A/B type using 24 routine blood parameters, achieving an AUC of 0.923 in influenza A detection, which is lower than the AUC

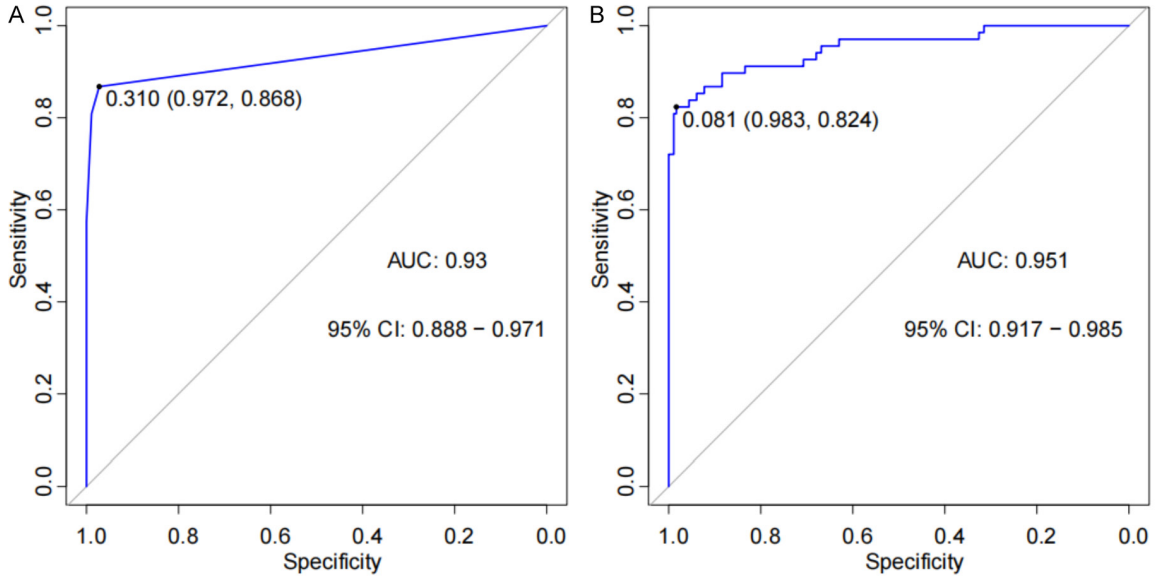

Figure 4. Receiver operator characteristic curve (ROC) for model prediction. A. Decision tree model; B. Random forest model. AUC: area under the curve; CI: confidence interval.

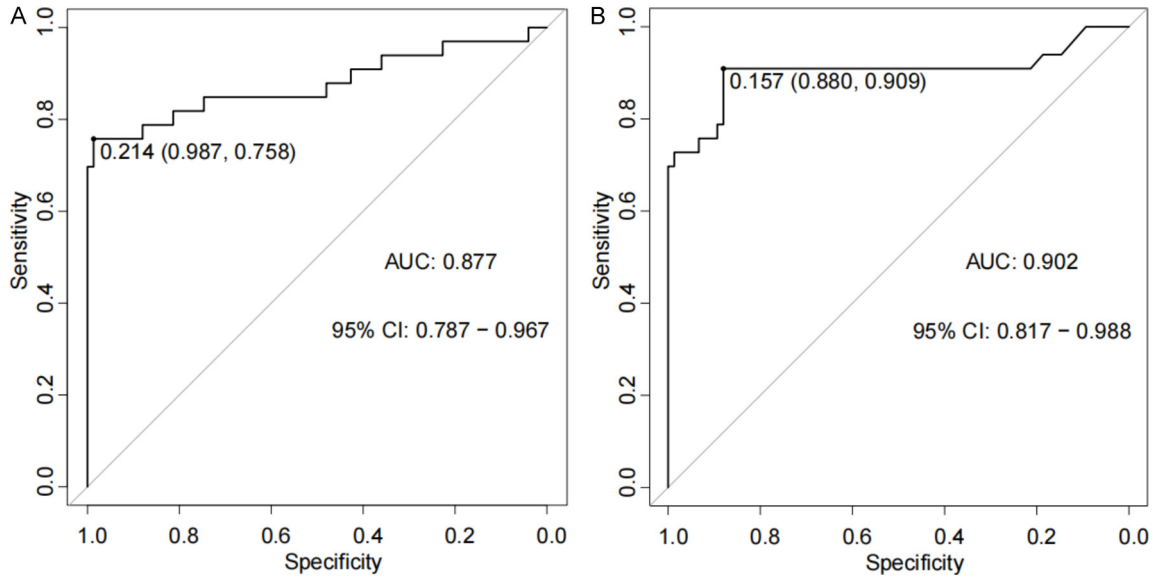

Figure 5. Receiver operator characteristic curve (ROC) for model validation. A. Decision tree model; B. Random forest model. AUC: area under the curve; CI: confidence interval.

of 0.951 of our random forest model in the modeling set, and comparable to the AUC of 0.878 of our logistic model. This difference stems from the fact that their study focused on virus typing, while our study is a specific prediction of secondary bacterial infection. In addition, we integrated multi-system indicators (such as coagulation function and myocardial enzymes) that can capture the pathophysiological complexity of bacterial co-infection.

Previous studies have mainly examined one or several inflammatory indicators, and there is a knowledge gap in the application of multidimensional routine clinical indicators to improve predictive accuracy. This study analyzed the data through multivariate logistic regression analysis and used a full range of routine indicators. The indicators used included age, routine blood parameters, inflammatory factors, myocardial enzymes, and coagulation function-

Table 6. Evaluation indicators of diagnostic efficacy for different models

Model	Data sets	Optimal cut-off	Sensitivity	Specificity	Accuracy	PPV	NPV	AUC
Logistic regression model	Modeling set	0.42	0.755	0.802	0.783	0.767	0.790	0.878
	Validation set	0.45	0.732	0.784	0.761	0.748	0.778	0.805
	External validation set	0.41	0.711	0.765	0.740	0.728	0.752	0.783
Decision tree model	Modeling set	0.48	0.868	0.972	0.900	0.950	0.920	0.930
	Validation set	0.46	0.758	0.987	0.850	0.960	0.820	0.877
	External validation set	0.43	0.771	0.803	0.794	0.786	0.792	0.823
Random forest model	Modeling set	0.50	0.824	0.983	0.900	0.970	0.880	0.951
	Validation set	0.49	0.909	0.880	0.838	0.824	0.831	0.902
	External validation set	0.47	0.802	0.832	0.823	0.812	0.822	0.852

PPV: positive predictive value; NPV: negative predictive value; AUC: area under the curve.

related indicators. After exploratory data analysis of clinical samples from 972 pneumonia patients, it was found that the synergistic effects of these different indicators have a stronger predictive ability for secondary pulmonary bacterial infection induced by influenza, thus expanding the dimensions of diagnostic indicators. In terms of model construction, logistic regression, as a classic statistical model, has certain diagnostic value in predicting influenza complications, but its generalization ability is limited [18]. It tends to underestimate the risk of rare but serious influenza-related complications because it essentially assumes a linear relationship between the predictor and the log odds of the outcome. This limitation makes it unable to capture the complex nonlinear interactions between clinical variables (e.g., inflammatory markers, comorbidities) in heterogeneous influenza patient populations [19]. A multicenter study by the Global Influenza Hospital Surveillance Network further confirmed that the predictive performance of logistic regression models established in a single-region population significantly decreases when applied to populations with different age structures and epidemiological backgrounds [20]. This study is the first to apply machine learning models (decision trees and random forests) to the auxiliary diagnosis of secondary pulmonary bacterial infection induced by influenza and compares their diagnostic efficacy with that of logistic regression models. This confirms the significant advantages of machine learning models in handling complex data relationships and generalizing across datasets, providing a novel technical approach for research in this field.

Exploration of potential mechanisms

The independent risk factors identified in this study are closely associated with the pathophysiological mechanisms of influenza-induced secondary pulmonary bacterial infection. Age is a key risk factor: older adults are more susceptible to such infections due to age-related decline in immune function and weakened respiratory mucosal defenses [21]. In elderly patients (≥65 years) after influenza infection, impaired mucociliary clearance, reduced alveolar macrophage phagocytosis, and dysregulated cytokine responses create a suitable microenvironment for the colonization of pathogens such as Streptococcus pneumoniae and Staphylococcus aureus [22]. A large-scale cohort study (1.3 million influenza-positive individuals) demonstrated that the incidence of secondary bacterial pneumonia in the 65-74 age group was 3.2 times higher than that in the 18-44 age group, and this age-related risk was independent of comorbidities [23]. Furthermore, advanced age (≥65 years) is an independent predictor of mortality related to secondary bacterial pneumonia, attributed to the synergistic effect of immunosenescence and epithelial barrier impairment [24]. WBC count is a direct marker of the body's immune response to pathogen infection. Bacterial infection typically triggers leukocytosis, a physiological process that plays a crucial role in pathogen clearance [25]. More specifically, influenza virus infection impairs the initial immune response. Subsequently, bacterial invasion triggers bone marrow hematopoiesis, promoting the release of granulocytes and their directed chemotaxis to the site of infection - an important compensatory strategy for the host against secondary bacterial infection [26].

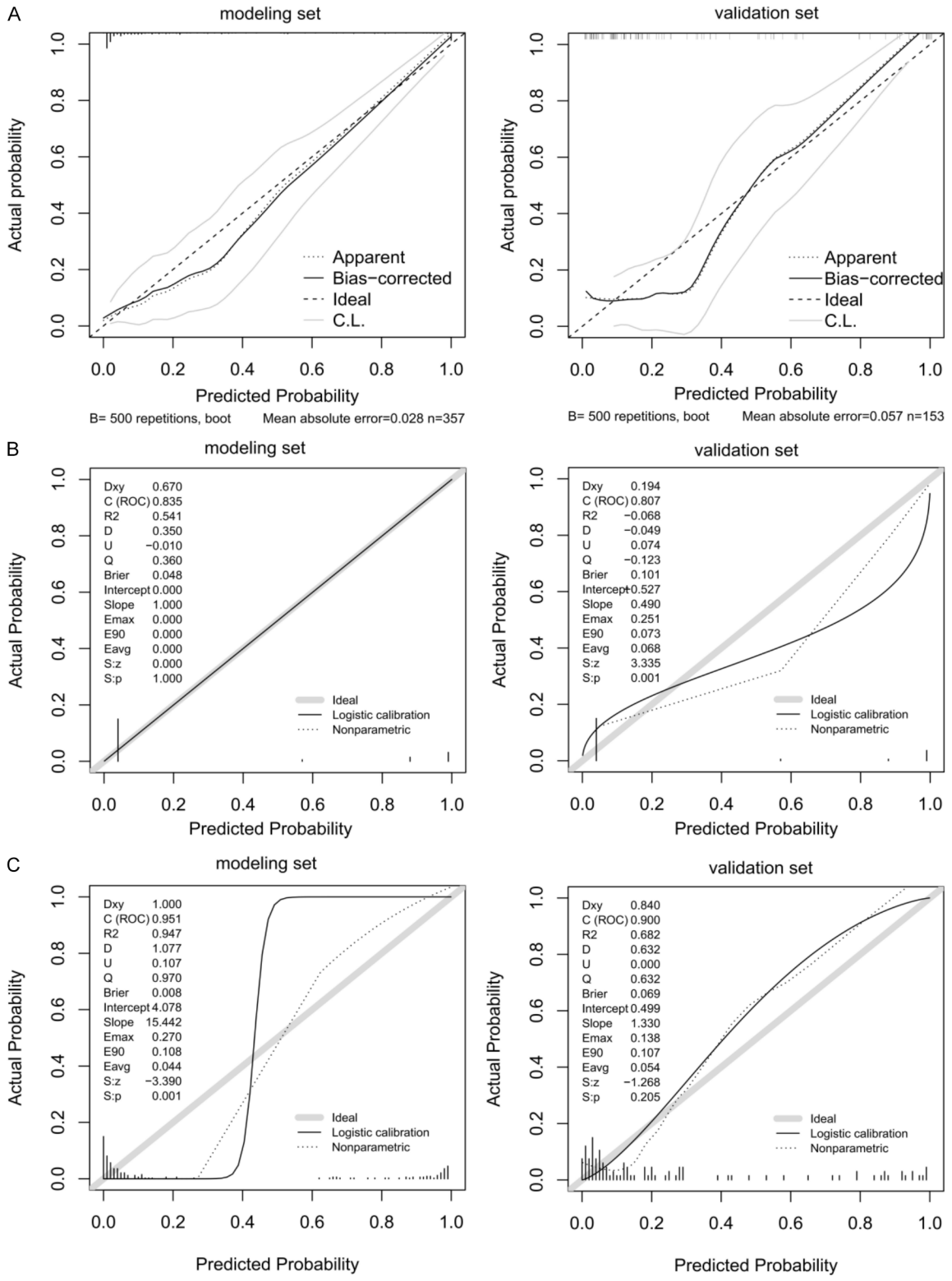

Figure 6. Calibration curves of the models. A. Logistic regression model; B. Decision tree model; C. Random forest model.

CRP and SAA are typical acute-phase reactants that increase during bacterial infection, primarily by activating the complement system and regulating immune effector functions, partici-

pating in the inflammatory cascade response to pathogens [27]. CRP can bind to antigens or bacterial cell wall components. Similar to other opsonins, opsonized bacteria are more readily phagocytosed by neutrophils and monocytes. SAA is thought to have a similar mechanism. Both proteins work together to clear pathogens. However, when the inflammatory system is uncontrolledly activated, it can lead to tissue damage [28]. Abnormalities in systems such as the heart and lungs are associated with abnormalities in CK-MB and D-dimer [29, 30]. Elevated CK-MB levels are generally associated with cardiomyocyte damage under inflammatory stress and are closely related to the severity of systemic inflammatory responses [31]. Elevated D-dimer levels indicate activation of the coagulation system, which is associated with enhanced microthrombus formation and fibrinolysis induced by pulmonary infection. These pathophysiological changes leading to impaired pulmonary blood flow and gas exchange can accelerate lesion formation [32, 33].

All these biomarkers may participate in the pathophysiology of influenza-associated secondary pulmonary bacterial infection through three related pathways (immunity, inflammation, and coagulation), further confirming the biological basis of the predictive power of our model.

Clinical importance

The diagnostic aid model developed in this study has significant clinical implications. When used for early diagnosis, this model, built upon routinely recorded clinical parameters, can generate early results, helping clinicians assess the likelihood of secondary pulmonary bacterial infection in influenza patients at their first visit, thus avoiding treatment delays. Importantly, it enables rapid case screening in primary care settings without specialized screening equipment. The correct application of this diagnostic tool may prevent the inappropriate use of antibiotics. Currently, the diagnostic specificity for influenza-associated secondary bacterial infections is weak, often leading to antibiotic overuse. By quantifying infection risk, this model provides an objective basis for the rational use of antibiotics and helps avoid the development of antimicrobial resistance, fully aligning with the principles of precision medicine. In addition, our machine learning model can be generalized across various clinical settings and patient populations. Validation on independent datasets indicates that our model contributes to improving the quality of diagnosis and treatment of influenza-associated secondary bacterial infections.

Limitations

This study has some limitations that must be addressed. First, the study sample may be subject to selection bias. Because all data were drawn from specific healthcare facilities, the patient population included in this study may not be as diverse as that in community-acquired or mixed-environment settings. This could limit the applicability of the model to heterogeneous influenza patient populations. Furthermore, since this study excluded patients with severe underlying conditions, the applicability of the findings may be limited to influenza patients without such comorbidities. Moreover, not all potential confounding factors were included in the study design. For example, it is known that patient comorbidities and influenza vaccination status affect susceptibility to post-influenza secondary bacterial infections, but these factors were not adequately considered in this study. Moreover, although our machine learning model shows promising performance, its internal algorithm is a “black box”: it is difficult to intuitively explain how our model makes decisions. This lack of transparency limits clinicians’ understanding and trust in the model’s predictions. Finally, the small sample size of the external validation cohort may affect a comprehensive assessment of the model’s stability and generalizability, especially in extreme clinical settings. Therefore, large-scale, multicenter studies are needed for further validation.

Conclusion

This study employed a multivariate logistic regression model to screen for independent risk factors for secondary pulmonary bacterial infection caused by influenza and age, WBC count, CRP, SAA, CK-MB and D-dimer levels. Using these identified indicators as core variables, we constructed logistic regression and machine learning diagnostic models. The results showed that, in terms of influenza diagnostic efficacy, the machine learning models (especially the random forest model) outperformed the classic logistic regression model,

exhibiting higher accuracy and better generalizability in assisting the diagnosis of influenza complicated by secondary pulmonary bacterial infection. This study offers new tools and insights for the prompt clinical identification and targeted intervention of influenza-associated secondary pulmonary bacterial infections.

Disclosure of conflict of interest

None.

Address correspondence to: Li-Zhi He, Department of Infectious Diseases, The Affiliated Yangming Hospital of Ningbo University (Yuyao People's Hospital), Yuyao 315400, Zhejiang, China. Tel: +86-0574-62619601; E-mail: he_lizhi1978@163.com

References

[1] Jasial S, Hu J, Miyao T, Hirama Y, Onishi S, Matsui R, Osaki K and Funatsu K. Screening and validation of odorants against influenza a virus using interpretable regression models. ACS Pharmacol Transl Sci 2022; 6: 139-150.

[2] Tomassini L, Ferretti F, Uvelli A, Fedeli D and Gualtieri G. Fatal viral and bacterial septicemia in a seventeen-year-old woman with immunodepressive influenza A H1N1: an autopsy case. Clin Ter 2024; 175: 95-100.

[3] Xie Y, Choi T and Al-Aly Z. Mortality in patients hospitalized for COVID-19 vs influenza in fall-winter 2023-2024. JAMA 2024; 331: 1963-1965.

[4] Bartley PS, Deshpande A, Yu PC, Klompas M, Haessler SD, Imrey PB, Zilberberg MD and Rothberg MB. Bacterial coinfection in influenza pneumonia: rates, pathogens, and outcomes. Infect Control Hosp Epidemiol 2022; 43: 212-217.

[5] Yang L, Li G, Yang J, Zhang T, Du J, Liu T, Zhang X, Han X, Li W, Ma L, Feng L and Yang W. Deep-learning model for influenza prediction from multisource heterogeneous data in a megacity: model development and evaluation. J Med Internet Res 2023; 25: e44238.

[6] Ogundimu EO, Altman DG and Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. J Clin Epidemiol 2016; 76: 175-182.

[7] Austin PC and Steyerberg EW. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. Stat Methods Med Res 2017; 26: 796-808.

[8] Alotaibi BS, Tantry BA, Bandy A, Ahmad R, Khursheed SQ, Ahmad A, Hakami MA and Shah NN. Simultaneous detection of influenza A/B, respiratory syncytial Virus, and SARS-CoV-2 in nasopharyngeal swabs by one-tube multiplex reverse transcription polymerase chain reaction. Trop Med Infect Dis 2023; 8: 326.

[9] Chen J and Xu F. Application of nanopore sequencing in the diagnosis and treatment of pulmonary infections. Mol Diagn Ther 2023; 27: 685-701.

[10] Ahmed WM, Fenn D, White IR, Dixon B, Nijsen TME, Knobel HH, Brinkman P, Van Oort PMP, Schultz MJ, Dark P, Goodacre R, Felton T, Bos LDJ and Fowler SJ; BreathDx Consortium. Microbial volatiles as diagnostic biomarkers of bacterial lung infection in mechanically ventilated patients. Clin Infect Dis 2023; 76: 1059-1066.

[11] Akbari H, Bushehri AP, Aqavil-Jahromi S, Eftekhari M, Esfehani KJ, Akhgar A and Jalili M. External validation of geriatric influenza death score: a prospective validation study. Aging Clin Exp Res 2026; 38: 57.

[12] Liu T, Zhang ZH, Zhou QH, Cheng QZ, Yang Y, Li JS, Zhang XM and Zhang JQ. MI-DenseCF-Net: deep learning-based multimodal diagnosis models for Aureus and Aspergillus pneumonia. Eur Radiol 2024; 34: 5066-5076.

[13] Zhou C, Jiang Y, Sun L, Li H, Liu X and Huang L. Secondary pulmonary infection and co-infection in elderly COVID-19 patients during the pandemics in a tertiary general hospital in Beijing, China. Front Microbiol 2023; 14: 1280026.

[14] Campani S, Talamonti M, Dall'Ara L, Coloretti I, Gatto I, Biagioni E, Tosi M, Meschiari M, Tonelli R, Clini E, Cossarizza A, Guaraldi G, Mussini C, Sarti M, Trenti T and Girardis M; MO-COVID-19 Working Group. The association of procalcitonin and C-reactive protein with bacterial infections acquired during intensive care unit stay in COVID-19 critically ill patients. Antibiotics (Basel) 2023; 12: 1536.

[15] Yang X, Zeng J, Yu X, Wang Z, Wang D, Zhou Q, Bai T and Xu Y. PCT, IL-6, and IL-10 facilitate early diagnosis and pathogen classifications in bloodstream infection. Ann Clin Microbiol Antimicrob 2023; 22: 103.

[16] Xiong S, Guo Y, Jiang L, Wang Y, Chen W, Bai Z, Zhu P, Qian J and Li L. A simple, rapid, and cost-effective model for predicting critical influenza a infection in children: a multicentre, retrospective cohort study. BMC Pediatr 2025; 25: 746.

[17] Hu W, Liu Y, Dong J, Peng X, Yang C, Wang H, Chen Y, Shi S and Li J. Evaluation of a machine learning model based on laboratory parameters for the prediction of influenza A and B in Chongqing, China: multicenter model development and validation study. J Med Internet Res 2025; 27: e67847.

[18] Tenforde MW, Noah KP, O'Halloran AC, Kirley PD, Hoover C, Alden NB, Armistead I, Meek J, Yousey-Hindes K, Openo KP, Witt LS, Monroe ML, Ryan PA, Falkowski A, Reeg L, Lynfield R,

McMahon M, Hancock EB, Hoffman MR, McGuire S, Spina NL, Felsen CB, Gaitan MA, Lung K, Shiltz E, Thomas A, Schaffner W, Talbot HK, Crossland MT, Price A, Masalovich S, Adams K, Holstein R, Sundaresan D, Uyeki TM, Reed C, Bozio CH and Garg S. Timing of influenza antiviral therapy and risk of death in adults hospitalized with influenza-associated pneumonia, influenza hospitalization surveillance network (FluSurv-NET), 2012-2019. Clin Infect Dis 2025; 80: 461-468.

[19] Cheong CW, Chen CL, Li CH, Seak CJ, Tseng HJ, Hsu KH, Ng CJ and Chien CY. Two-stage prediction model for in-hospital mortality of patients with influenza infection. BMC Infect Dis 2021; 21: 451.

[20] Cohen LE, Hansen CL, Andrew MK, McNeil SA, Vanhems P, Kyncl J, Domingo JD, Zhang T, Dbaibo G, Laguna-Torres VA, Draganescu A, Baumeister E, Gomez D, Raboni SM, Giamberardino HIG, Nunes MC, Burtseva E, Sominina A, Medić S, Coulibaly D, Salah AB, Otieno NA, Koul PA, Unal S, Tanriover MD, Mazur M, Bresee J, Viboud C and Chaves SS. Predictors of severity of influenza-related hospitalizations: results from the global influenza hospital surveillance network (GIHSN). J Infect Dis 2024; 229: 999-1009.

[21] Sanches Santos Rizzo Zuttion M, Parimon T, Bora SA, Yao C, Lagree K, Gao CA, Wunderink RG, Kitsios GD, Morris A, Zhang Y, McVerry BJ, Modes ME, Marchevsky AM, Stripp BR, Soto CM, Wang Y, Merene K, Cho S, Victor BL, Vujkovic-Cvijin I, Gupta S, Cassel SL, Sutterwala FS, Devkota S, Underhill DM and Chen P. Antibiotic use during influenza infection augments lung eosinophils that impair immunity against secondary bacterial pneumonia. J Clin Invest 2024; 134: e180986.

[22] Mandell LA, Wunderink RG, Anzueto A, Bartlett JG, Campbell GD, Dean NC, Dowell SF, File TM Jr, Musher DM, Niederman MS, Torres A and Whitney CG; Infectious Diseases Society of America; American Thoracic Society. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. Clin Infect Dis 2007; 44 Suppl 2: S27-72.

[23] Shirata M, Ito I, Jo T, Iwao T, Oi I, Hamao N, Nishioka K, Yamana H, Nagase T, Yasunaga H and Hirai T. Factors associated with the development of bacterial pneumonia related to seasonal influenza virus infection: a study using a large-scale health insurance claim database. Open Forum Infect Dis 2023; 10: ofad222.

[24] Yi G, de Kraker MEA, Buetti N, Zhong X, Li J, Yuan Z, Zhu W, Zhou J and Zhou H. Risk factors for in-hospital mortality and secondary bacterial pneumonia among hospitalized adult patients with community-acquired influenza: a large retrospective cohort study. Antimicrob Resist Infect Control 2023; 12: 25.

[25] Moser D, Feuerecker M, Biere K, Han B, Hoerl M, Schelling G, Kaufmann I, Choukér A and Woehrle T. SARS-CoV-2 pneumonia and bacterial pneumonia patients differ in a second hit immune response model. Sci Rep 2022; 12: 15485.

[26] Su G, Chen Y, Li X and Shao JW. Virus versus host: influenza A virus circumvents the immune responses. Front Microbiol 2024; 15: 1394510.

[27] Atosuo J, Karhuvaara O, Suominen E, Virtanen J, Vilén L and Nuutila J. The role of gamma globulin, complement component 1q, factor B, properdin, body temperature, C-reactive protein and serum amyloid alpha to the activity and the function of the human complement system and its pathways. J Immunol Methods 2024; 531: 113709.

[28] Peng Y, Fang Y, Li Z, Liu C and Zhang W. Saa3 promotes pro-inflammatory macrophage differentiation and contributes to sepsis-induced AKI. Int Immunopharmacol 2024; 127: 111417.

[29] Scarsi E, Dorighi U, Adriano E, Grandis M and Balestrino M. Low-dose creatine supplementation may be effective in early-stage statin myopathy: a preliminary study. J Clin Med 2024; 13: 7194.

[30] Suárez-Cuenca JA, Flores-Zaleta JF, Corona-Rojas LA, Guzmán-Rullán P, Camacho-Barajas LA, Aguilera-Ontiveros U, Melchor-López A, González-Mora A, Guzmán-Ramírez PM and Rodríguez-Solis J. Tomographic features of lung damage associate with D-Dimer levels and further clinical outcome in patients with acute respiratory distress syndrome due to COVID-19. BMC Pulm Med 2025; 25: 65.

[31] Ilhan I, Asci H, Tepebasi MY, Imeci OB, Sevuk MA, Temel EN and Ozmen O. Selenium exerts protective effects on inflammatory cardiovascular damage: molecular aspects via SIRT1/p53 and Cyt-c/Cas-3 pathways. Mol Biol Rep 2023; 50: 1627-1637.

[32] Gupta V, Acharya S and Keerti A. Common coagulopathies associated with COVID-19 patients. Cureus 2023; 15: e38067.

[33] Bhalla M, Herring S, Lenhard A, Wheeler JR, Aswad F, Klumpp K, Rebo J, Wang Y, Wilhelmsen K, Fortney K and Bou Ghanem EN. The prostaglandin D2 antagonist asapiprant ameliorates clinical severity in young hosts infected with invasive *Streptococcus pneumoniae*. Infect Immun 2024; 92: e0052223.