Infection and Drug Resistance

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

A Predictive Model for Pulmonary Aspergillosis in ICU Patients: A Multicenter Retrospective Cohort Study

Yujing Li[1,2,*], Xindie Ren[3,*], Qianqian Wang[4,*], Songying Shen[2,*], Yihao Li[1,2], Xinling Qian[2], Yufei Tang[2], Jinguang Jia[2], Hao Zhang[2], Junjie Ding[2], Yinsen Song[1], Sisen Zhang[1], Shengfeng Wang[5], Yinghe Xu[6], Yongpo Jiang[6], Xuwei He[7], Muhua Dai[8], Lin Zhong[9], Yonghui Xiong[10], Yujie Pan[11], Mingqiang Wang[12], Huanzhang Shao[12], Hongliu Cai[3], Lingtong Huang[3], Hongyu Wang[1,2]

[1]Department of Critical Care Medicine, The Fifth Clinical Medical College of Henan University of Chinese Medicine, Zhengzhou, Henan Province, People's Republic of China; [2]Department of Critical Care Medicine, People's Hospital of Henan University of Chinese Medicine/People's Hospital of Zhengzhou, Zhengzhou, Henan Province, People's Republic of China; [3]Department of Critical Care Medicine, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang Province, People's Republic of China; [4]Department of Critical Care Medicine, The First Hospital of Jiaxing, Jiaxing, Zhejiang Province, People's Republic of China; [5]Department of Critical Care Medicine, The Second Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan Province, People's Republic of China; [6]Department of Critical Care Medicine, Taizhou Hospital of Zhejiang Province affiliated with Wenzhou Medical University, Taizhou, Zhejiang Province, People's Republic of China; [7]Department of Critical Care Medicine, Lishui People's Hospital, Lishui, Zhejiang Province, People's Republic of China; [8]Department of Critical Care Medicine, Tongde Hospital of Zhejiang Province, Hangzhou, Zhejiang Province, People's Republic of China; [9]Department of Critical Care Medicine, The First People's Hospital of Pinghu, Pinghu, Zhejiang Province, People's Republic of China; [10]Department of Critical Care Medicine, Lanxi Hospital of Traditional Chinese Medicine, Lanxi, Zhejiang Province, People's Republic of China; [11]Department of Critical Care Medicine, Wenzhou Central Hospital, Wenzhou, Zhejiang Province, People's Republic of China; [12]Department of Critical Care Medicine, Henan Key Laboratory for Critical Care Medicine, Zhengzhou Key Laboratory for Critical Care Medicine, Henan Provincial People's Hospital; Zhengzhou University People's Hospital, Henan University People's Hospital, Zhengzhou, Henan Province, People's Republic of China

*These authors contributed equally to this work

Correspondence: Lingtong Huang, Department of Critical Care Medicine, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, People's Republic of China, Email lingtonghuang@zju.edu.cn; Hongyu Wang, Department of Critical Care Medicine, People's Hospital of Henan University of Chinese Medicine/People's Hospital of Zhengzhou, Zhengzhou, Henan Province, People's Republic of China, Email hongyu260@163.com

Background: Several predictive models for invasive pulmonary aspergillosis (IPA) based on clinical characteristics have been reported. Nevertheless, the significance of other concurrently detected microorganisms in IPA patients is equally noteworthy. This study aimed to develop a risk prediction model for IPA by integrating clinical and microbiological characteristics.

Methods: This retrospective study was conducted in adult intensive care units (ICUs) of 17 medical centers in China. Clinical data were collected from patients with severe pneumonia who underwent clinical metagenomics of bronchoalveolar lavage fluid between January 1, 2019, and June 30, 2023. Subsequently, patients were randomly assigned to training and validation cohorts in a 7:3 ratio. In the training cohort, potential influencing factors were identified through univariate analysis, clinical practice, and existing literature, and a risk prediction model was constructed using multivariate logistic regression analysis. The performance of this model was then assessed and validated in the validation cohort.

Results: Out of 1737 patients initially included in the study, 898 were ultimately analyzed, of which 100 (11%) were diagnosed with IPA. The risk prediction model for IPA, incorporating microbiological characteristics, identified six independent risk factors, namely age, immunosuppression, chronic kidney disease, connective tissue disease, liver failure, and cytomegalovirus positivity. The model demonstrated a superior discriminative ability, with area under the curve (AUC) values of 0.791 and 0.792 in the training and validation cohorts, respectively. Sensitivity and specificity reached 73.1% and 74.9%, respectively, and the model demonstrated good calibration.

Conclusion: This study developed a novel risk prediction model for IPA incorporating microbiological characteristics based on clinical metagenomics. The model exhibited good discriminative ability and calibration.

Keywords: CAP, community-acquired pneumonia, IPA, invasive pulmonary aspergillosis, prediction model, microbiological characteristics, clinical metagenomics

Received: 25 August 2024
Accepted: 12 January 2025
Published: 23 January 2025

Background

Aspergillus is one of the most common causes of fungal infections in humans.[1] According to statistics, in 2017, there were over 1.8 million cases of invasive fungal infections globally, with approximately 250,000 cases of invasive pulmonary aspergillosis (IPA).[2] The mortality rate for invasive pulmonary aspergillosis (IPA) is extremely high, ranging from 45% to 90%.[3,4] Lack of understanding of host factors in critically ill patients, non-specific pulmonary imaging, and the ambiguous clinical significance of positive *Aspergillus* cultures from airway secretions lead to delayed diagnosis and treatment of IPA in the intensive care unit (ICU), thereby contributing to elevated mortality rates.[5–7] At present, clinical diagnostic algorithms and prediction models for IPA have been developed.[8–11] However, these models have traditionally been based solely on clinical features. Meanwhile, for patients with IPA, the detection of other concurrently present microorganisms, such as *Pseudomonas aeruginosa, Klebsiella pneumoniae*, respiratory syncytial virus, Epstein-Barr virus (EBV), and cytomegalovirus, is as important.[12] Nevertheless, there is currently a lack of IPA prediction models that also incorporate microbiological features.

Recently, clinical metagenomics has been widely used in diagnosing infectious diseases, particularly for the non-targeted diagnosis of specific, unknown, or mixed pathogens.[13,14] Its distinct advantage lies in unbiased sampling, which enables the broad identification of known and unexpected pathogens and even the discovery of new organisms, thus providing a comprehensive overview of pathogens in a given sample.[15,16] Therefore, this study aimed to use clinical metagenomics methods to identify microorganisms associated with IPA. Additionally, it sought to construct an IPA risk prediction model that combines clinical and microbiological characteristics and to evaluate its predictive performance.

Methods

Patient Enrollment

This retrospective cohort study was conducted in the adult ICUs of 17 medical centers in China. We collected clinical data of all patients admitted to the ICU between January 1, 2019, and June 30, 2023. Inclusion criteria were: 1. age ≥ 18 years; 2. diagnosed with community-acquired pneumonia (CAP); 3. undergoing invasive mechanical ventilation; and 4. receiving commercial metagenomic next-generation sequencing (mNGS) of bronchoalveolar lavage fluid. Exclusion criteria were: 1. loss to follow-up within 28 days of ICU admission; 2. diagnosis or clinical suspicion of IPA before ICU admission. This study received approval from the ethics committees of all participating hospitals. This study was conducted in accordance with the declaration of Helsinki. As a retrospective study, informed consent was waived.

Definitions and Data Collection

The diagnosis of CAP was based on the official clinical practice guidelines of the American Thoracic Society and the Infectious Diseases Society of America (IDSA).[17] Specifically, it was defined as community onset with chest imaging showing new patchy infiltrates, lobar or segmental consolidation, ground-glass opacities or interstitial changes, with or without pleural effusion, and any pneumonia-related clinical manifestation. The diagnostic criteria for IPA, we use the 2021 EORTC/MSG criteria from IDSA, which are more suitable for ICU patients. This standard classifies IPA into proven and probable.[18] Based on the IDAS criteria, our study considers clinical metagenomic testing showing positive for *Aspergillus* to be consistent with mycological evidence. Since biopsy was not feasible in the ICU, all patients were classified as probable cases. The 100 probable IPA cases in this study were diagnosed based on meeting at least one mycological evidence, one clinical feature, and one host factor. Immunosuppression was defined as (1) neutropenia (a neutrophil count < 0.5 × 10^9/L within 10 days of admission); (2) use of immunosuppressive drugs, including tacrolimus, cyclosporine, mycophenolate mofetil, or monoclonal antibodies (eg, rituximab) within 30 days before mNGS testing; and (3) history of acquired immunodeficiency syndrome, hematologic malignancies, or transplantation.[19]

Relevant data were independently collected from patients' electronic medical records by an experienced team of clinicians. For the included patients, data obtained included gender, age, comorbidities, immunosuppressive status, laboratory test results, galactomannan test results, sputum fungal culture results, and clinical metagenomics results. Disease severity was assessed using the sequential organ failure assessment (SOFA) score at ICU admission, with organ dysfunction defined as a score of ≥ 2 points. All ICU centers collected cases according to unified standards. The clinical metagenomics laboratories were accredited by either the College of American Pathologists or the external quality assessment programs of the National Health Commission of China.[20,21]

Model Construction

Factors believed to be associated with IPA were included in a multivariate logistic regression model based on previous literature and clinical expertise.[11] These variables were then screened using the forward selection method, and those with a p-value < 0.05 were considered independent risk factors for IPA and included in the analysis. Two risk prediction models were constructed based on the results of univariate analysis and multivariate logistic regression: one without microbiological characteristics (Model 1) and one with microbiological characteristics (Model 2). The models constructed by multivariate logistic regression were visualized using nomograms. Model performance was evaluated using data from the training cohort. Discriminatory ability of the models was assessed by calculating the area under the receiver operating characteristic curve (AUC). Model calibration was performed using the Hosmer–Lemeshow (HL) goodness-of-fit test, with a p-value > 0.05 indicating an acceptable fit. Calibration curves were plotted, and decision curves were used to evaluate the clinical benefit of the models. Predictive abilities of the two models were compared using net reclassification improvement (NRI). An NRI > 0 indicated improved predictive ability of the new model compared to the old model, NRI < 0 indicated a decline, and NRI = 0 indicated no significant difference. The models were internally validated using the validation cohort data through AUC, HL test, and decision curve analysis (DCA).

Statistical Analysis

Continuous data were first tested for normality. Normally distributed data were expressed as mean ± standard deviation ($\bar{x} \pm s$) and compared using independent sample *t*-tests. Non-normally distributed data were expressed as median (interquartile range) [M (QL, QU)] and compared using the Wilcoxon rank-sum test. Categorical data were expressed as frequencies (percentages) and compared using the χ^2-test, with the continuity correction χ^2-test employed when expected values were < 5. All tests were two-sided, and a p value less than 0.05 was considered statistically significant. There were no missing data in this dataset. During the construction of the IPA risk prediction models, the "pROC" and "ggplot2" packages in R were used to plot receiver operating characteristic (ROC) curves and evaluate model discrimination. The "rmst" package and Bootstrap method, with 1000 repeated samples, were employed to plot calibration curves and test model fit. Decision curves were plotted using the "dcurves" and "rmda" packages to evaluate the clinical benefit of the models. The "nricens" package was applied to calculate NRI to compare the predictive abilities of the two models. Statistical analysis was performed using SPSS 23.0 software (SPSS Inc.) and R Statistics software 4.4.0.

Results

Basic Characteristics of Enrolled Patients

Out of 1897 patients screened, 898 met the inclusion criteria and were included in this study (Figure 1). According to the revised IDSA diagnostic criteria, 100 patients were diagnosed with probable IPA (100/898, 11%), of which 70 cases (70/100, 70%) exhibited *Aspergillus* positivity from clinical metagenomics, and 38 cases (38/100, 38%) showed *Aspergillus* positivity from sputum fungal culture. Among the 100 patients, 74 died within 28 days, resulting in a 28-day ICU mortality rate of 74% for the IPA group.

Compared to the non-IPA group, patients in the IPA group had higher incidences of concurrent myocardial infarction (12.0% vs 4.9%, p = 0.004), CKD (26.0% vs 12.4%, p < 0.001), hematologic malignancies (13.0% vs 3.0%, p < 0.001), CTD (15.0% vs 3.8%, p < 0.001), and history of transplantation (15.0% vs 4.6%, p < 0.001). Additionally, the IPA group exhibited lower lymphocyte counts [0.46 (0.21–0.77) vs 0.54 (0.31–0.90), p = 0.014] and CRP levels [69.82 (31.54–-152.21) vs 98.19 (43.02–171.35), p = 0.031]. Clinical metagenomics indicated that patients in the IPA group were more likely to develop co-infections with *Pneumocystis spp*. (23.0% vs 9.4%, p < 0.001), EBV (31.0% vs 15.7%, p < 0.001), and cytomegalovirus (33.0% vs 16.3%, p< 0.001), whereas patients in the control group were more likely to experience co-infections with *Klebsiella spp*. (16.0% vs 29.4%, p = 0.005). The IPA group also had significantly shorter hospital stays [15 (7–23) vs 18 (10–31), p < 0.001], shorter ICU stays [9 (6–15) vs 13 (8–22.25), p < 0.001], and a higher 28-day mortality rate (74.0% vs 47.0%, p < 0.001) compared to the non-IPA group (Table 1). In this study, the 898 included patients were randomly divided into a training cohort (n = 637) and a validation cohort (n = 261) in a 7:3 ratio. There were no notable differences between the two cohorts in terms of gender, age, comorbidities, immunosuppressive status,

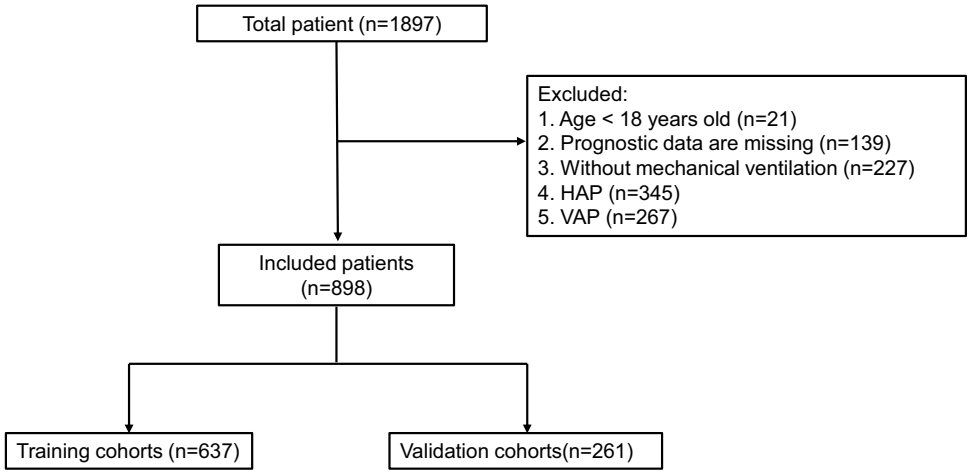

Figure 1 Flow chart.

laboratory test results, SOFA scores, degree of organ dysfunction, duration of hospital stay, clinical metagenomics results, duration of ICU stay, and 28-day mortality rate ($p > 0.05$) (Table S1).

Comparison Between IPA Group and Control Group Based on Univariate Analysis

In the training cohort, patients diagnosed with probable IPA were designated as the IPA group (n = 67), while the remaining patients were designated as the control group (n = 570). Compared to the control group, the IPA group had higher incidences of CKD (28.4% vs 12.5%, $p < 0.001$), hematologic malignancies (13.4% vs 3.7%, $p = 0.001$), CTD (16.4% vs 3.9%, $p < 0.001$), and a history of transplantation (16.4% vs 5.1%, $p = 0.001$). Additionally, a higher proportion of patients in the IPA group were in an immunosuppressive state (64.2% vs 21.8%, $p < 0.001$), with lower CRP levels [68.04 (27.05–130.00) vs 105.00 (42.94–180.83), $p = 0.011$] (Table 2).

Table 1 Baseline Characteristics and Outcomes in Full Population

Variables	IPA (n=100)	Non-IPA (n=798)	P value
Age, years, median (IQR)	68.00 (62.00–76.00)	68.00 (56.00–77.00)	0.373
Male gender, n (%)	71 (71.0)	549 (68.8)	0.653
Comorbidities n (%)			
Diabetes mellitus	28 (28.0)	216 (27.1)	0.843
Myocardial infarction	12 (12.0)	39 (4.9)	0.004
Liver disease	10 (10.0)	51 (6.4)	0.176
Chronic kidney disease	26 (26.0)	99 (12.4)	<0.001
Solid tumor	19 (19.0)	107 (13.4)	0.129
Hematologic malignancy	13 (13.0)	24 (3.0)	<0.001
Connective tissue disease	15 (15.0)	30 (3.8)	<0.001
Transplantation	15 (15.0)	37 (4.6)	<0.001
Cerebrovascular disease	16 (16.0)	138 (17.3)	0.746
Immunosuppression, n (%)	63 (63.0)	171 (21.4)	<0.001
Laboratory indicators, median (IQR)			
White blood cell (10^9/L)	12.25 (6.19–15.60)	11.23 (6.99–16.34)	0.755
Lymphocyte (10^9/L)	0.46 (0.21–0.77)	0.54 (0.31–0.90)	0.014
Neutrophil (10^9/L)	11.16 (5.40–14.65)	9.86 (5.77–14.53)	0.695
C reactive protein (mg/L)	69.82 (31.54–152.21)	98.19 (43.02–171.35)	0.031
Procalcitonin (ng/mL)	0.78 (0.24–3.52)	1.14 (0.27–8.18)	0.143

(*Continued*)

Table 1 (Continued).

Variables	IPA (n=100)	Non-IPA (n=798)	P value
SOFA score at transfer to ICU, median (IQR)	8.00 (5.00–10.00)	7.00 (5.00–10.00)	0.229
Organ dysfunction, n (%)			
Respiratory	88 (88.0)	715 (89.6)	0.624
Coagulation	39 (39.0)	210 (26.3)	0.008
Liver	22 (22.0)	102 (12.8)	0.012
Cardiovascular	52 (52.0)	394 (49.4)	0.620
Neurological	23 (23.0)	251 (31.5)	0.084
Kidney	29 (29.0)	178 (22.3)	0.134
Time from ICU admission to mNGS testing, days, median (IQR)	3.00 (2.00–4.00)	3.00 (2.00–4.00)	0.501
Clinical metagenomics results, n (%)			
Acinetobacter spp.	20 (20.0)	222 (27.8)	0.097
Klebsiella spp.	16 (16.0)	235 (29.4)	0.005
Pseudomonas spp.	12 (12.0)	111 (13.9)	0.601
Stenotrophomonas spp.	12 (12.0)	110 (13.8)	0.623
Enterococcus spp.	18 (18.0)	117 (14.7)	0.379
Burkholderia spp.	8 (8.0)	54 (6.8)	0.647
Staphylococcus spp.	5 (5.0)	79 (9.9)	0.113
Corynebacterium spp.	4 (4.0)	43 (5.4)	0.557
Escherichia spp.	2 (2.0)	34 (4.3)	0.415
Streptococcus spp.	5 (5.0)	79 (9.9)	0.113
Haemophilus spp.	3 (3.0)	34 (4.3)	0.741
Elizabethkingia spp.	1 (1.0)	26 (3.3)	0.349
Achromobacter spp.	2 (2.0)	22 (2.8)	0.910
Enterobacter spp.	1 (1.0)	16 (2.0)	0.760
Candida spp.	30 (30.0)	238 (29.8)	0.971
Pneumocystis spp.	23 (23.0)	75 (9.4)	<0.001
Aspergillus spp.	70 (70.0)	77 (9.6)	<0.001
Torque teno virus	9 (9.0)	74 (9.3)	0.929
Nakaseomyces spp.	4 (4.0)	50 (6.3)	0.369
Serratia spp.	2 (2.0)	17 (2.1)	1.000
HSV-1	30 (30.0)	186 (23.3)	0.140
EBV	31 (31.0)	125 (15.7)	<0.001
CMV	33 (33.0)	130 (16.3)	<0.001
HHV-7	1 (1.0)	32 (4.0)	0.220
HHV-6b	1 (1.0)	13 (1.6)	0.960
Duration of mechanical ventilation within 28 days, days, median (IQR)	8.00 (4.00–14.00)	9.00 (5.00–16.00)	0.068
Hospital stays, day, median (IQR)	15.00 (7.00–23.00)	18.00 (10.00–31.00)	0.001
ICU stay, day, median (IQR)	9.00 (6.00–15.00)	13.00 (8.00–22.25)	<0.001
28-day mortality in ICU, n (%)	74 (74.0)	375 (47.0)	<0.001

Note: Data are presented as median (interquartile range), n (%).
Abbreviations: CMV, Cytomegalovirus; EBV, Epstein-Barr virus; HSV, Herpes simplex virus; HHV, Human herpes virus; IQR, interquartile range; ICU, Intensive Care Unit; mNGS, Metagenomic next-generation sequencing; SOFA, Sequential Organ Failure Assessment;

Model Construction

In this study, variables considered significant in previous literature were included in the multivariate logistic regression model for analysis. Five variables were found to be independent risk factors for IPA: age, immunosuppression, CKD, CTD, and liver failure (Table 3). Subsequently, based on the multivariate logistic regression analysis, an IPA risk prediction model was constructed based on clinical characteristics (Model 1). The regression system of each self-variable was used as the weight system, the logistic regression equation was established, and the final fitting risk prediction model was:

Table 2 Comparisons Between the IPA and Control Groups for Univariate Analysis

Variables	IPA (n=67)	Non-IPA (n=570)	P value
Age, years, median (IQR)	66.00 (62.00–74.00)	56.00 (68.00–77.00)	0.680
Male gender, n (%)	51 (76.1)	392 (68.8)	0.216
Comorbidities, n (%)			
Diabetes mellitus	21 (31.3)	163 (28.6)	0.639
Myocardial infarction	8 (11.9)	30 (5.3)	0.056
Liver disease	5 (7.5)	39 (6.8)	1.000
Chronic kidney disease	19 (28.4)	71 (12.5)	<0.001
Solid tumor	10 (14.9)	74 (13.0)	0.657
Hematologic malignancy	9 (13.4)	21 (3.7)	0.001
Connective tissue disease	11 (16.4)	22 (3.9)	<0.001
Transplantation	11 (16.4)	29 (5.1)	0.001
Cerebrovascular disease	10 (14.9)	107 (18.8)	0.442
Immunosuppression, n (%)	43 (64.2)	124 (21.8)	<0.001
Laboratory indicators, median (IQR)			
White blood cell (10^9/L)	11.67 (6.60–14.57)	11.31 (6.96–16.92)	0.883
Lymphocyte (10^9/L)	0.52 (0.25–0.88)	0.56 (0.30–0.95)	0.244
Neutrophil (10^9/L)	10.67 (5.90–13.27)	9.90 (5.73–14.79)	0.856
C reactive protein (mg/L)	68.04 (27.05–130.00)	105.00 (42.94–180.83)	0.011
Procalcitonin (ng/mL)	0.56 (0.22–2.98)	1.16 (0.27–9.94)	0.053
SOFA score at transfer to ICU, median (IQR)	7.00 (5.00–10.00)	7.00 (5.00–10.00)	0.788
Organ dysfunction, n (%)			
Respiratory	56 (83.6)	503 (88.2)	0.271
Coagulation	24 (35.8)	156 (27.4)	0.146
Liver	15 (22.4)	71 (12.5)	0.024
Cardiovascular	31 (46.3)	228 (50.5)	0.510
Neurological	13 (19.4)	162 (28.4)	0.118
Kidney	22 (32.8)	129 (22.6)	0.063

Note: Data are presented as median (interquartile range), n (%).
Abbreviations: IQR, interquartile range; ICU, Intensive Care Unit.

Table 3 Multivariate Logistic Regression Model-1

Variables	β	P value	OR	95% CI
Age	0.024	0.018	1.024	1.004–1.044
Immunosuppression	1.709	<0.001	5.521	3.107–9.811
CKD	0.813	0.012	2.255	1.193–4.263
CTD	0.979	0.025	2.661	1.131–6.260
Liver	0.861	0.015	2.366	1.181–4.741
Constant	−4.828	<0.001	0.008	

Abbreviations: CKD, Chronic kidney disease; CTD, Connective tissue disease; Liver, Liver dysfunction.

$$\mathrm{Ln}[P/1-P] = -4.828 + 0.024 \times \mathrm{Age} + 1.709 \times \mathrm{Immunosuppression} + 0.813 \times \mathrm{CKD} + 0.979 \times \mathrm{CTD} + 0.861 \times \mathrm{Liver}$$

The nomogram based on this model is shown in Figure 2A.

Clinical metagenomics analysis indicated that, compared to the control group, the IPA group was more likely to develop co-infections with *Klebsiella spp., Pneumocystis spp*., EBV, and cytomegalovirus (Table 4). A second risk prediction model incorporating microbial characteristics (Model 2) was then established. In Model 2,

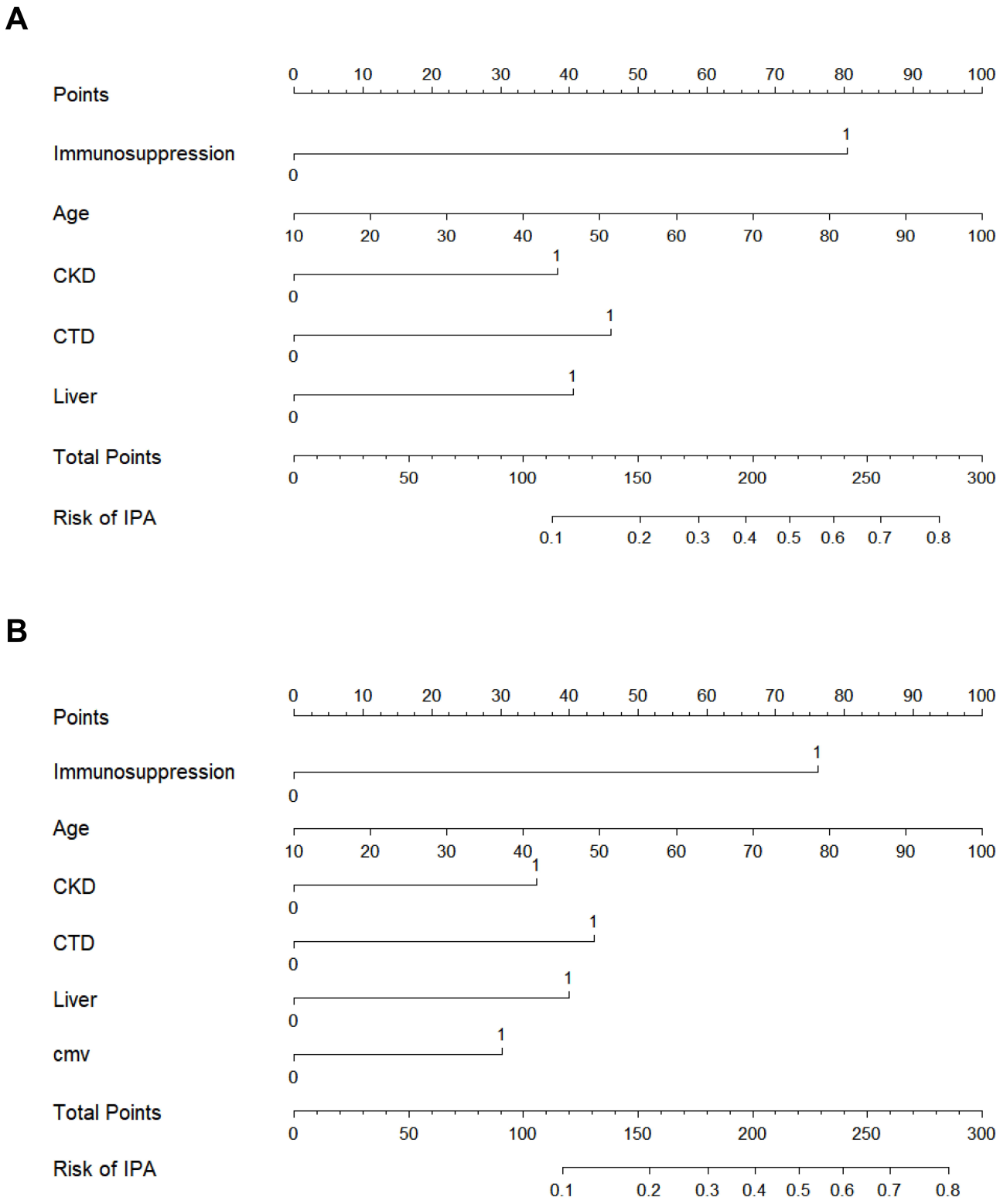

Figure 2 (**A**). Nomogram prediction model-1; (**B**). Nomogram prediction model-2.

six factors were identified as independent risk factors for IPA, namely age, immunosuppression, CKD, CTD, liver failure, and cytomegalovirus positivity (Table 5). The regression system of each self-variable was used as the weight system, the logistic regression equation was established, and the final fitting risk prediction model was:

https://doi.org/10.2147/IDR.S493019

Table 4 Microbial Analysis of IPA and Control Groups in the Training Cohort

Variables	IPA (n=67)	Non-IPA (n=570)	P value
Acinetobacter spp.	15 (12.4)	163 (28.6)	0.284
Klebsiella spp.	9 (13.4)	173 (30.4)	0.004
Pseudomonas spp.	11 (16.4)	81 (14.2)	0.627
Stenotrophomonas spp.	6 (9.0)	81 (14.2)	0.236
Enterococcus spp.	15 (22.4)	86 (15.1)	0.122
Burkholderia spp.	7 (10.4)	36 (6.3)	0.309
Staphylococcus spp.	3 (4.5)	59 (10.4)	0.125
Corynebacterium spp.	3 (4.5)	30 (5.3)	1.000
Escherichia spp.	2 (3.0)	24 (4.2)	0.878
Streptococcus spp.	5 (7.5)	62 (10.9)	0.389
Haemophilus spp.	3 (4.5)	24 (4.2)	1.000
Elizabethkingia spp.	1 (1.5)	18 (3.2)	0.705
Achromobacter spp.	2 (3.0)	18 (3.2)	1.000
Enterobacter spp.	1 (1.5)	14 (2.5)	0.947
Candida spp.	20 (29.9)	164 (28.8)	0.854
Pneumocystis spp.	15 (22.4)	59 (10.4)	0.004
Aspergillus spp.	47 (70.1)	55 (9.6)	<0.001
Torque teno virus	4 (6.0)	57 (10.0)	0.289
Nakaseomyces spp.	2 (3.0)	33 (5.8)	0.503
Serratia spp.	2 (3.0)	15 (2.6)	1.000
HSV 1	20 (29.9)	122 (21.4)	0.116
EBV	21 (31.3)	90 (15.8)	0.001
CMV	23 (34.3)	88 (15.4)	<0.001
HHV 7	1 (1.5)	21 (3.7)	0.565
HHV 6b	–	10 (1.8)	0.566

Note: Data are presented as n (%).
Abbreviations: CMV, Cytomegalovirus; EBV, Epstein-Barr virus; HSV, Herpes simplex virus; HHV, Human herpes virus.

Table 5 Multivariate Logistic Regression Model-2

Variables	β	P value	OR	95% CI
Age	0.024	0.019	1.024	1.004–1.044
Immunosuppression	1.620	<0.001	5.055	2.822–9.058
CKD	0.749	0.023	2.116	1.107–4.045
CTD	0.928	0.036	2.530	1.061–6.030
Liver	0.851	0.018	2.341	1.158–4.733
CMV	0.641	0.041	1.899	1.027–3.512
Constant	−4.920	<0.001	0.007	

Abbreviations: CKD, Chronic kidney disease; CTD, Connective tissue disease; Liver, Liver dysfunction; CMV, Cytomegalovirus.

$$\mathrm{Ln}[P/1-P] = 4.9200.024 \times \mathrm{Age} + 1.620 \times \mathrm{Immunosuppression} + 0.749 \times \mathrm{CKD} + 0.928 \times \mathrm{CTD} + 0.851 \times \mathrm{Liver} + 0.641 \times \mathrm{Cytomegalovirus}$$

The nomogram based on this model is shown in Figure 2B.

Model Evaluation and Validation

ROC curves were plotted using training cohort data. Model 1 was observed to have an AUC of 0.782 (0.725–0.840), and Model 2 had an AUC of 0.791 (0.735–0.847). Both models demonstrated good discriminative ability, with overall AUC

values higher than those of any individual parameter within the models (Figure 3A). In the validation cohort, the AUCs of Model 1 and Model 2 were 0.787 (0.703–0.871) and 0.792 (0.710–0.874), respectively, again exceeding the AUC values of any individual parameter (Figure 3B). Calibration curves based on the training cohort revealed that the mean absolute error (MAE) for Models 1 and 2 were 0.012 and 0.011, respectively, and closely matched the ideal curve. The HL test indicated p-values of 0.263 for Model 1 and 0.323 for Model 2, suggesting good model fit (Figure 4A and B). In the validation cohort, the MAEs for Models 1 and 2 were 0.017 and 0.02, respectively, with HL test p-values of 0.252 and 0.083 (Figure 4C and D).

Sensitivity analysis of the models was conducted based on the Youden index. In the training cohort, Model 1 exhibited a sensitivity of 0.806 and a specificity of 0.698 at an optimal cut-off value of 0.083. In contrast, Model 2 had a sensitivity of 0.731 and a specificity of 0.749 at an optimal cut-off value of 0.103 (Table 6). The positive predictive values for Models 1 and 2 were 0.239 (0.183–0.295) and 0.255 (0.194–0.317), respectively, while their negative predictive values were 0.968 (0.951–0.985) and 0.960 (0.941–0.978), respectively. Comparison of the two models yielded an NRI = 0.035 > 0 (Table 6), indicating improved predictive capability of Model 2 over Model 1.

DCA suggested that in the validation cohort, Model 1 had a higher net benefit than the "All" and "None" lines between thresholds of 8% and 38%, whereas Model 2 showed a higher net benefit between thresholds of 8% and 50%. The net benefit area for Model 2 was larger than that for Model 1, indicating the superiority of the former. In the training cohort, Model 1 exhibited a higher net benefit than the "All" and "None" lines between thresholds of 5% and 37%, while Model 2 showed a higher net benefit than the "All" and "None" lines between thresholds of 8% and 35% (Figure 5).

Discussion

There is an increasing incidence of IPA, especially among critically ill hospitalized patients.[22] Early identification and treatment of IPA are closely linked to reduced mortality rates. Although traditional culture methods are considered the gold standard for diagnosing IPA due to their accuracy in identifying strains, they are time-consuming and yield low positive rates, making them unsuitable for early clinical diagnosis.[23] Non-culture methods such as histopathological examination, while significant, cannot distinguish species and involve invasive sampling processes often limited by the patient's condition, thus restricting their clinical application.[24] Therefore, it is necessary to develop clinical prediction models that can help predict the likelihood of IPA at an earlier stage. In 2020, Huang developed a predictive scoring system for influenza-associated aspergillosis (IAA) called Asper-PreSS.[11] In 2023, Massart developed a prediction model

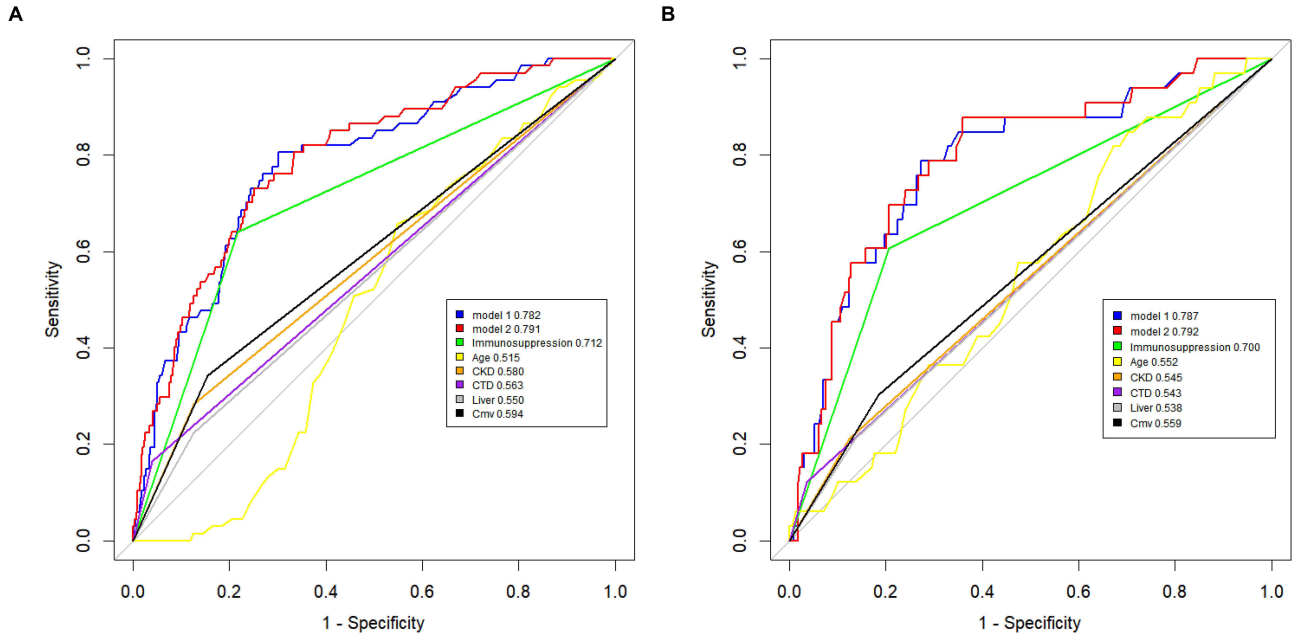

Figure 3 (**A**). The ROC curves of Model 1 and Model 2 in the training cohort; (**B**). The ROC curves of Model 1 and Model 2 in the validation cohort.

https://doi.org/10.2147/IDR.S493019

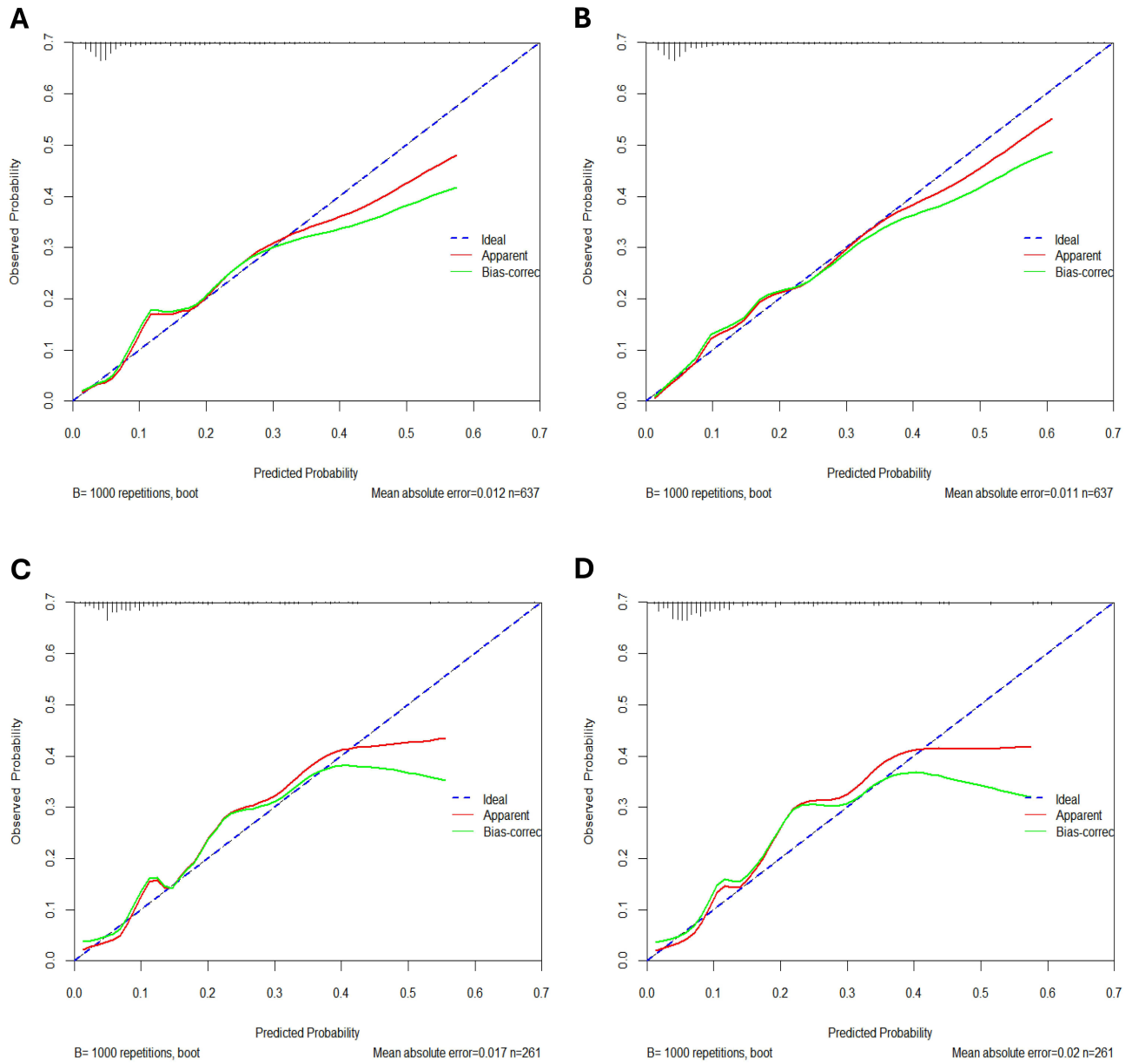

Figure 4 (**A**). Calibration curve of model 1 in the training cohort; (**B**). Calibration curve of model 2 in the training cohort; (**C**). Calibration curve of model 1 in the validation cohort; (**D**). Calibration curve of model 2 in the validation cohort.

for IPA in patients with ventilator-associated pneumonia.[8] However, to our knowledge, previous predictive models for IPA have only focused on the clinical characteristics of IPA. This study is the first to combine clinical characteristics and microbiota to construct a predictive model for IPA.

While numerous IPA prediction models based on clinical characteristics have been proposed, this study is the first to incorporate microbial characteristics alongside clinical features.

This study evaluated the incidence of IPA among 898 patients with CAP who underwent invasive mechanical ventilation. Using readily available variables from early ICU admission or pre-admission stages, 12 influencing factors were identified through univariate analysis of clinical characteristics, clinical practice, and previous literature. Subsequently, a risk prediction model based on clinical characteristics (Model 1) was constructed and validated using logistic regression analysis. The model comprised five predictive factors, namely age, immunosuppression, CKD, CTD,

Table 6 Sensitivity Analysis of the Two Models

	Training cohort		Validation Cohort	
	Model 1	**Model 2**	**Model 1**	**Model 2**
Cutoff	0.083	0.103	0.106	0.088
Sensitivity	0.806 (0.711–0.901)	0.731 (0.625–0.837)	0.788 (0.648–0.927)	0.879 (0.767–0.990)
Specificity	0.698 (0.661–0.736)	0.749 (0.714–0.785)	0.728 (0.670–0.786)	0.640 (0.578–0.703)
PPV	0.239 (0.183–0.295)	0.255 (0.194–0.317)	0.295 (0.200–0.391)	0.261 (0.180–0.343)
NPV	0.968 (0.951–0.985)	0.960 (0.941–0.978)	0.960 (0.930–0.989)	0.973 (0.948–0.999)
PLR	2.671 (2.250–3.171)	2.915 (2.380–3.571)	2.897 (2.197–3.820)	2.443 (1.972–3.028)
NLR	0.278 (0.170–0.454)	0.359 (0.241–0.534)	0.291 (0.150–0.565)	0.189 (0.075–0.447)
AUC	0.782 (0.725–0.840)	0.791 (0.735–0.847)	0.787 (0.702–0.872)	0.792 (0.709–0.875)
NRI	–	0.035	–	0.009

Note: Sensitivity, Specificity, PPV, NPV, PLR, NLR, and AUC are presented as point estimates (95% Confidence Interval).
Abbreviations: PPV, positive predictive value; NPV, Negative predictive value; PLR, positive likelihood ratio; NLR, negative likelihood ratio; AUC, area under the curve; NRI, Net reclassification index.

and liver failure, aligning with previous research findings.[8,11,25–27] An observational study revealed that patients with invasive *Aspergillus* infections typically presented with more underlying diseases and were often immunosuppressed.[7]

After establishing Model 1, it was evaluated and validated. The AUCs of the model in the training and validation cohorts were 0.782 (0.725–0.840) and 0.787 (0.703–0.871), respectively, indicating a high discriminative ability. Additionally, the p-values of the HL test for both cohorts were > 0.05, suggesting a good model fit. Sensitivity analysis of the model revealed a sensitivity of 80.6% and specificity of 69.8% at an optimal cut-off value of 0.083. Furthermore, DCA for the training and validation cohorts demonstrated that the model exhibited significant clinical benefits within a certain range.

Based on clinical characteristics, microbial features were further analyzed and compared between the IPA and non-IPA groups. It was observed that the IPA group was more likely to develop co-infections with *Pneumocystis spp*., EBV, and cytomegalovirus, while the control group was more likely to experience co-infection with *Klebsiella spp*. Subsequently, the microbial characteristics were combined with clinical features to construct a risk prediction model (Model 2). This model consisted of six predictive factors: age, immunosuppression, CKD, CTD, liver failure, and cytomegalovirus positivity, consistent with the findings of previous research.[8,11,25–28] IPA has been reported to be closely

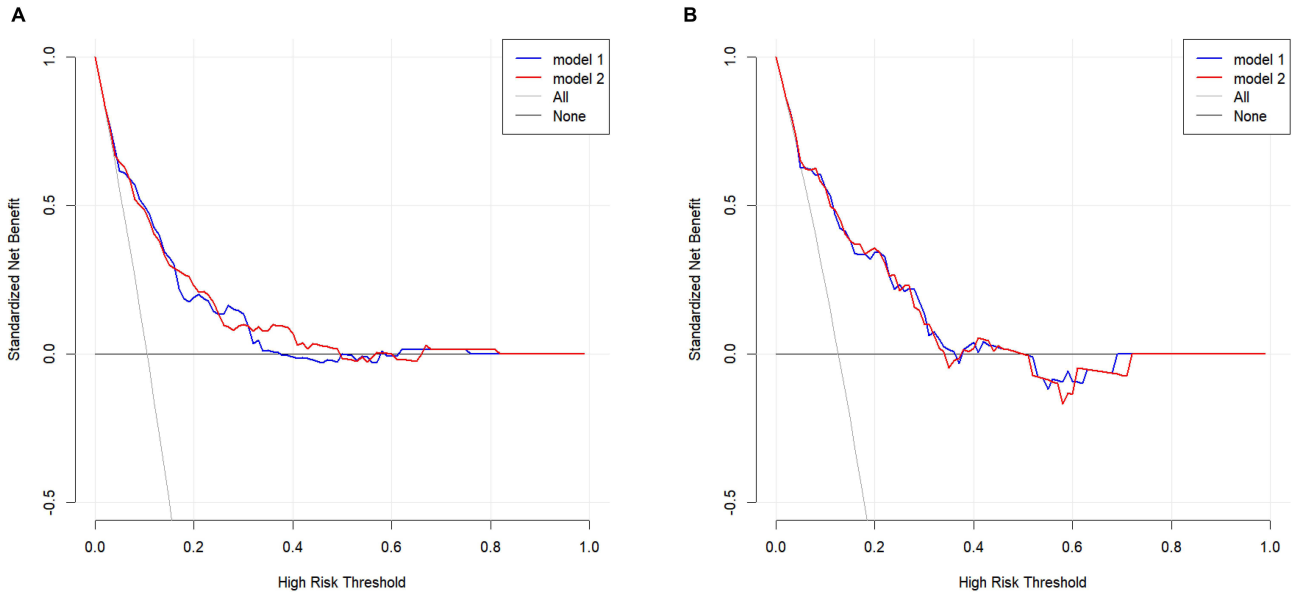

Figure 5 (**A**). Decision curves for model 1 and model 2 in the training cohort; (**B**). Decision curves for model 1 and model 2 in the validation cohort.

https://doi.org/10.2147/IDR.S493019

associated with cytomegalovirus, as a retrospective study has demonstrated that cytomegalovirus is an independent risk factor for IPA.[28] Following its establishment, the model was evaluated, validated, and compared against Model 1 to explore its clinical utility. The AUCs of Model 2 in the training and validation cohorts were 0.791 (0.735–0.847) and 0.792 (0.710–0.874), respectively, indicating a high discriminative ability in both cohorts. Although Model 2 had a higher AUC compared to Model 1, the p-value for the AUCs between the two models was 0.35, indicating no statistically significant difference. This suggests that the clinical benefit of Model 2 may be limited, possibly due to the insufficient sample size. Additionally, the p-values of the HL test for both the training and validation cohorts were > 0.05, indicating superior model fit and robustness. Compared to Model 1, Model 2 produced a better HL test result in the training cohort but performed worse in the validation cohort, likely due to the smaller sample size. Sensitivity analysis suggested that at an optimal cut-off value of 0.083, the sensitivity was 73.1% and the specificity was 74.9%. Compared to Model 1, the sensitivity of Model 2 decreased while the specificity increased. Furthermore, the DCA results in the training cohort showed that Model 2 exhibited better clinical utility than Model 1. The NRI was calculated to compare the predictive capabilities of the two models. In both the training and validation cohorts, the NRI was > 0, indicating that Model 2 had better predictive ability for events than Model 1.

This study has some distinct advantages. Firstly, to the authors' knowledge, this is the first study to propose an IPA risk prediction model in the ICU that integrates microbial characteristics with clinical features, and the proposed model demonstrated high discriminative ability and appropriate calibration. Secondly, through clinical metagenomics, we have identified the specific microorganisms involved in the occurrence of IPA. In subsequent clinical practice, we can replace clinical metagenomics with PCR detection of these microorganisms. Thirdly, our statistical methods are quite comprehensive. Using univariate analysis and logistic multivariable regression methods to construct predictive models in the training cohort. This method is suitable for binary outcome variables. Subsequently, use a nomogram to visually present the constructed model. Using ROC, calibration curves, and decision curves to evaluate the model. In addition, we also validated the constructed model in the validation cohort using internal validation methods. Finally, a quantitative comparison of Model 1 and Model 2 was conducted. Making our research more reliable.

However, the study also has some limitations. First, this study is a retrospective study with some missing data. Second, clinical metagenomics only reported coexisting species, without providing absolute quantification. Third, this study did not undergo external validation, which may result in overfitting of the model. A well-designed prospective cohort can serve as prospective validation of this model[29]。

Conclusion

This was the first study that combined microbiological features with clinical characteristics to construct a clinically applicable IPA risk prediction model. The model exhibited superior discriminative ability and calibration, as well as high sensitivity and specificity. These findings suggest that incorporating microbial characteristics can significantly improve the early identification and management of IPA in critically ill patients, potentially leading to timely and targeted therapeutic interventions and ultimately reducing mortality rates. Future studies should focus on external validation of this model and explore its applicability in different clinical settings to enhance its generalizability and reliability.

Abbreviations

AUC, Area under the curve; CAP, Community-acquired pneumonia; CKD, Chronic kidney disease; CRP, C-reactive protein; CTD, Connective tissue disease; DCA, Decision curve analysis; EBV, Epstein-Barr virus; HL – Hosmer, Lemeshow; ICU, Intensive care unit; IDSA, Infectious Diseases Society of America; IPA, Invasive pulmonary aspergillosis; MAE, Mean absolute error; mNGS, Metagenomic next-generation sequencing; NRI, Net reclassification improvement; ROC, Receiver operating characteristic; SOFA, Sequential organ failure assessment.

Data Sharing Statement

The datasets used during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

This study was conducted with approval from the Ethics Committee of Zhengzhou People's Hospital. This study was conducted in accordance with the principles of the Declaration of Helsinki. The need for informed consent was waived by the ethics committees as this was a retrospective study that only involved the review of de-identified patient data. All patient information was anonymized and handled with strict confidentiality throughout the study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the Zhejiang Provincial Natural Science Fund (LTGY24H190001), National Natural Science Foundation of China (82202356).

Disclosure

The authors declare that they have no competing interests in this work.

References

1. Wei X, Huang X, Gu S, et al. Landscape of fungal detection in the lungs of patients with severe pneumonia in the ICU, a multicenter study based on clinical metagenomics. *J Infect*. 2024;89:106195. doi:10.1016/j.jinf.2024.106195
2. Bongomin F, Gago S, Oladele RO, Denning DW. Global and multi-national prevalence of fungal diseases-estimate precision. *J Fungi*. 2017;3:57. doi:10.3390/jof3040057
3. Verweij PE, Brüggemann RJM, Azoulay E, et al. Taskforce report on the diagnosis and clinical management of COVID-19 associated pulmonary aspergillosis. *Intensive Care Med*. 2021;47:819–834. doi:10.1007/s00134-021-06449-4
4. Verweij PE, Rijnders BJA, Brüggemann RJM, et al. Review of influenza-associated pulmonary aspergillosis in ICU patients and proposal for a case definition: an expert opinion. *Intensive Care Med*. 2020;46:1524–1535. doi:10.1007/s00134-020-06091-6
5. Schauwvlieghe AFAD, Rijnders BJA, Philips N, et al. Invasive aspergillosis in patients admitted to the intensive care unit with severe influenza: a retrospective cohort study. *Lancet Respir Med*. 2018;6:782–792. doi:10.1016/S2213-2600(18)30274-1
6. Cadena J, Thompson GR, Patterson TF. Aspergillosis: epidemiology, diagnosis, and treatment. *Infect Dis Clin North Am*. 2021;35:415–434. doi:10.1016/j.idc.2021.03.008
7. Taccone FS, Van den Abeele A-M, Bulpa P, et al. Epidemiology of invasive aspergillosis in critically ill patients: clinical presentation, underlying conditions, and outcomes. *Crit Care*. 2015;19:7. doi:10.1186/s13054-014-0722-7
8. Massart N, Plainfosse E, Benameur Y, et al. Prediction of pulmonary aspergillosis in patients with ventilator-associated pneumonia. *Ann Intensive Care*. 2023;13:109. doi:10.1186/s13613-023-01199-6
9. Hamam J, Navellou J-C, Bellanger A-P, et al. New clinical algorithm including fungal biomarkers to better diagnose probable invasive pulmonary aspergillosis in ICU. *Ann Intensive Care*. 2021;11:41. doi:10.1186/s13613-021-00827-3
10. Blot SI, Taccone FS, Van den Abeele A-M, et al. A clinical algorithm to diagnose invasive pulmonary aspergillosis in critically ill patients. *Am J Respir Crit Care Med*. 2012;186:56–64. doi:10.1164/rccm.201111-1978OC
11. Huang L, Zhai T, Hua L, Zhan Q. Early identification of patients with severe influenza-associated aspergillosis (IAA) in the intensive care unit——an IAA prediction score system (Asper-PreSS). *J Infect*. 2020;81:639–646. doi:10.1016/j.jinf.2020.07.036
12. Hérivaux A, Willis JR, Mercier T, et al. Lung microbiota predict invasive pulmonary aspergillosis and its outcome in immunocompromised patients. *Thorax*. 2022;77:283–291. doi:10.1136/thoraxjnl-2020-216179
13. Huang L, Zhang X, Fang X. Case report: Epstein-Barr virus encephalitis complicated with brain stem hemorrhage in an immune-competent adult. *Front Immunol*. 2021;12:618830. doi:10.3389/fimmu.2021.618830
14. Zhang X, Wang J, Huang X, et al. Case report: parvovirus B19 infection complicated by hemophagocytic lymphohistiocytosis in a heart-lung transplant patient. *Front Immunol*. 2023;14:1099468. doi:10.3389/fimmu.2023.1099468
15. Gwinn M, MacCannell D, Armstrong GL. Next-generation sequencing of infectious pathogens. *JAMA*. 2019;321:893–894. doi:10.1001/jama.2018.21669
16. Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet*. 2019;20:341–355. doi:10.1038/s41576-019-0113-7
17. Olson G, Davis AM. Diagnosis and treatment of adults with community-acquired pneumonia. *JAMA*. 2020;323:885–886. doi:10.1001/jama.2019.21118
18. Bassetti M, Azoulay E, Kullberg B-J, et al. EORTC/MSGERC definitions of invasive fungal diseases: summary of activities of the intensive care unit working group. *Clin Infect Dis*. 2021;72:S121–7. doi:10.1093/cid/ciaa1751
19. Huang L, Zhang X, Pang L, et al. Viral reactivation in the lungs of patients with severe pneumonia is associated with increased mortality, a multicenter, retrospective study. *J Med Virol*. 2023;95(e28337). doi:10.1002/jmv.28337

20. Xu J, Zhong L, Shao H, et al. Incidence and clinical features of HHV-7 detection in lower respiratory tract in patients with severe pneumonia: a multicenter, retrospective study. *Crit Care*. 2023;27:248. doi:10.1186/s13054-023-04530-6
21. Jiang Y, Huang X, Zhou H, et al. Clinical characteristics and prognosis of patients with severe pneumonia with pneumocystis jirovecii colonization: a multicenter, retrospective study. *Chest*. 2024;S0012-3692(24):4841.
22. Koulenti D, Papathanakos G, Blot S. Invasive pulmonary aspergillosis in the ICU: tale of a broadening risk profile. *Curr Opin Crit Care*. 2023;29:463–469. doi:10.1097/MCC.0000000000001070
23. Bassetti M, Bouza E. Invasive mould infections in the ICU setting: complexities and solutions. *J Antimicrob Chemother*. 2017;72:i39–47. doi:10.1093/jac/dkx032
24. Yeo SF, Wong B. Current status of nonculture methods for diagnosis of invasive fungal infections. *Clin Microbiol Rev*. 2002;15:465–484. doi:10.1128/CMR.15.3.465-484.2002
25. Li W, Chen G, Lin F, et al. A score for predicting invasive pulmonary aspergillosis in immunocompetent critically ill patients. *Eur J Clin Invest*. 2023;53:e13985. doi:10.1111/eci.13985
26. Contou D, Dorison M, Rosman J, et al. Aspergillus-positive lower respiratory tract samples in patients with the acute respiratory distress syndrome: a 10-year retrospective study. *Ann Intensive Care*. 2016;6:52. doi:10.1186/s13613-016-0156-2
27. Shekhova E, Salazar F, Da Silva Dantas A, et al. Age difference of patients with and without invasive aspergillosis: a systematic review and meta-analysis. *BMC Infect Dis*. 2024;24:220. doi:10.1186/s12879-024-09109-2
28. Pardo E, Lemiale V, Mokart D, et al. Invasive pulmonary aspergillosis in critically ill patients with hematological malignancies. *Intensive Care Med*. 2019;45:1732–1741. doi:10.1007/s00134-019-05789-6
29. Wei X, Guo L, Cai H, et al. MASS cohort: multicenter, longitudinal, and prospective study of the role of microbiome in severe pneumonia and host susceptibility. *iMeta [Internet]*. 2024 [cited 2024 Jul 27]. Available from: https://webofscience.clarivate.cn/api/gateway?GWVersion=2&SrcAuth=DOISource&SrcApp=UA&KeyAID=10.1002%2Fimt2.218&DestApp=DOI&SrcAppSID=USW2EC0FDFvldsp2z8WDOVM5qe4SJ&SrcJTitle=IMETA&DestDOIRegistrantName=Wiley+%28John+Wiley+%26+Sons%29. Accessed January 16, 2025.