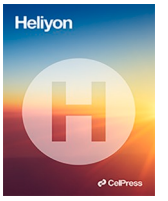
Heliyon 10 (2024) e37367

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Development and validation of machine learning-based prediction model for severe pneumonia: A multicenter cohort study

Zailin Yang [a,1], Shuang Chen [a,1], Xinyi Tang [a,b], Jiao Wang [c], Ling Liu [d], Weibo Hu [e], Yulin Huang [f], Jian'e Hu [g], Xiangju Xing [h], Yakun Zhang [a,b], Jun Li [a], Haike Lei [a,**], Yao Liu [a,*]

[a] *Department of Hematology-Oncology, Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing University Cancer Hospital, Chongqing, 400030, China*
[b] *School of Medicine Chongqing University, Chongqing, 400044, China*
[c] *Department of Medical Laboratory, Chongqing General Hospital, Chongqing, 401121, China*
[d] *Department of Medical Laboratory, the People's Hospital of Chongqing Liangjiang New Area, Chongqing, 401121, China*
[e] *Department of Medical Laboratory, the People's Hospital of Rongchang District, Chongqing, 402460, China*
[f] *Department of Medical Laboratory, the People's Hospital of Kaizhou District, Chongqing, 405499, China*
[g] *Department of Medical Laboratory, the Three Gorges Hospital Affiliated of Chongqing University, Chongqing, 404000, China*
[h] *Department of Respiratory Medicine, the Third Affiliated Hospital of Chongqing Medical University, Chongqing, 401120, China*

ARTICLE INFO

Keywords:
Severe pneumonia
Inflammatory factors
Machine learning
Prediction
Validation

ABSTRACT

Severe pneumonia (SP) is a prevalent respiratory ailment characterized by high mortality and poor prognosis. Current scoring systems for pneumonia are not only time-consuming but also exhibit limitations in early SP prediction. To address this gap, this study aimed to develop a machine-learning model using inflammatory markers from peripheral blood for early prediction of SP. A total of 204 pneumonia patients from seven medical centers were studied, with 143 (68 SP cases) in the training cohort and 61 (32 SP cases) in the test cohort. Clinical characteristics and laboratory test results were collected at diagnosis. Various models including Logistic Regression, Random Forest, Naïve Bayes, XGBoost, Support Vector Machine, and Decision Tree were built and evaluated. Seven predictors—age, sex, WBC count, T-lymphocyte count, NLR, CRP, TNF-α, IL-4/IFN-γ ratio, IL-6/IL-10 ratio—were selected through LASSO regression and clinical insight. The XGBoost model, exhibiting best performance, achieved an AUC of 0.901 (95 % *CI*: 0.827 to 0.985) in the test cohort, with an accuracy of 0.803, sensitivity of 0.844, specificity of 0.759, and F1_score of 0.818. Indeed, SHAP analysis emphasized the significance of elevated WBC counts, older age, and elevated CRP as the top predictors. The use of inflammatory biomarkers in this concise predictive model shows significant potential for the rapid assessment of SP risk, thereby facilitating timely preventive interventions.

* Corresponding author.
** Corresponding author.
E-mail addresses: tohaike@163.com (H. Lei), liuyao77@cqu.edu.cn (Y. Liu).
[1] Contribution to this work has been equally, with each sharing first authorship.

https://doi.org/10.1016/j.heliyon.2024.e37367
Received 11 April 2024; Received in revised form 30 July 2024; Accepted 2 September 2024
Available online 3 September 2024

1. Introduction

Pneumonia stands as a prevalent acute respiratory ailment, particularly afflicting children and older adults owing to their weaker immune functions [1]. Severe pneumonia (SP) is the most critical manifestation of this condition and presents with life-threatening complications including respiratory failure, systemic inflammatory response, septic shock, or multiple organ dysfunction [2]. The majority of SP patients are diagnosed at a moderate or severe stage, with survival rates ranging from 24 % to 50 %, highlighting the urgency of early prediction and prompt medical intervention to reduce mortality [3]. Moreover, a potential global outbreak of pneumonia similar to COVID-19 could lead to a dramatic rise in pneumonia cases, overwhelming healthcare systems and amplifying the demand for effective pneumonia prediction tools. Currently, the CURB-65 and Pneumonia Severity Index (PSI) are widely used tools for assessing pneumonia severity. However, these scores require time-consuming evaluations and have limitations in the early prediction of SP due to the evaluating criteria being mainly dependent on clinical features [4,5]. Recent evidence suggests that patients with severe or critically ill pneumonia typically exhibit extreme systemic inflammation and poorer prognosis [6].

Importantly, blood-related laboratory tests for inflammation have been developed as significant complementary methods for evaluating severity. Among the routine inflammatory biomarkers used in clinical practice, such as C-reactive protein (CRP), procalcitonin, white blood cell (WBC) count, and neutrophil-to-lymphocyte ratio (NLR), there is growing interest in their potential as prognostic indicators [7]. In addition, novel inflammatory biomarkers, particularly cytokine profiles, have garnered attention for their ability to reflect the balance between proinflammatory and anti-inflammatory factors [8,9]. Multiple studies have highlighted the detrimental role of elevated inflammatory cytokines in disease progression and their direct link to pulmonary damage in patients with SP [10,11]. Cytokines are widespread in diverse body fluids, notably peripheral blood serum, bronchoalveolar lavage fluid, and cerebrospinal fluid [12]. Previous findings have highlighted interleukin-6 (IL-6), IL-10, IL-17, and tumor necrosis factor α (TNF-α) as potential biomarkers for early diagnosis and indicators of disease severity in various contexts [10,13,14]. Furthermore, other inflammatory markers like T lymphocyte subsets count, $CD14^{+}HLA\text{-}DR^{-}$ monocyte , and neutrophil CD64 (nCD64) index, have been implicated in conditions ranging from COVID-19 [7] to Tuberculosis [15] and Chagas Disease [16]. Nevertheless, the definitive early predictive capacity of those biomarkers specifically for SP remains elusive.

In this study, we developed a model to investigate the clinical value of combined inflammation-associated biomarkers for the rapid prediction of SP in patients with community-acquired pneumonia (CAP). By harnessing the power of machine learning (ML), a branch of artificial intelligence, we integrated multiple clinical and laboratory datasets to enhance prediction accuracy using a promising approach [17]. We systematically screened potential inflammatory biomarkers predictive of SP and constructed a predictive model by analyzing data from both SP and non-severe pneumonia (nSP) groups. Furthermore, we rigorously evaluated the performance of our models using key metrics, including the receiver operating characteristic (ROC) curve, accuracy, sensitivity, specificity, and F1-score. To assess the clinical feasibility of our model, we performed decision curve analysis (DCA) and calibration curve analysis. This comprehensive strategy endeavors to deepen our comprehension of biomarkers linked to disease severity, thereby enabling earlier diagnosis and guiding targeted treatment strategies.

2. Materials and methods

2.1. Study patients

All patients were enrolled during the period between April 2021 and August 2022 from seven medical centers in Chongqing. The seven centers are (1) Chongqing People's Hospital; (2) The First People's Hospital of Chongqing Liang Jiang New Area; (3) Rong Chang District People's Hospital of Chongqing; (4) Kai Zhou District People's Hospital of Chongqing; (5) Three Gorges Central Hospital Affiliated to Chongqing University; (6) The Third Affiliated Hospital of Chongqing Medical University; (7) Chongqing university cancer hospital. The included patients were classified as newly diagnosed SP and nSP by the 2019 American Thoracic Society and Infectious Disease Society of America (ATS/IDSA) guidelines [18]. This study was conducted with the approval of the Ethics Committee of Chongqing University Cancer Hospital and met the Helsinki Declaration.

2.2. Inclusion and exclusion criteria

Patients who met the following criteria were enrolled: (1) patients with age $\geq$18 years; (2) diagnosis of SP and nSP according to the 2019 ATS/IDSA guidelines [18]. The exclusion criteria included: (1) recent medication use including hormones, granulocyte colony-stimulating factor, cyclophosphamide, cyclosporin, interferon, and tumor necrosis factor-alpha antagonists within the past 28 days; (2) human immunodeficiency virus infection, hematologic tumor, organ transplantation; (3) recent partial lung resection or major surgical interventions; (4) severe neuropsychiatric disorders; (5) death within 24 h of admission; (6) missing data.

2.3. Diagnostic criteria

The diagnostic criteria of the guidelines for nSP include (1) community-onset illness; (2) Clinical manifestations related to pneumonia: (a). New onset of cough, sputum production, or worsening of preexisting respiratory disease symptoms, with or without purulent sputum, chest pain, dyspnea, and hemoptysis; (b). Fever; (c). Signs of lung consolidation and/or moist rales on auscultation; (d). Peripheral white blood cell count $>10 \times 10^9$/L or $<4 \times 10^9$/L, with or without a left shift in the nuclear segmentation; and (3) Chest imaging shows new patchy infiltrates, lobar or segmental consolidation, ground-glass opacities, or interstitial changes, with or

without pleural effusion. A diagnosis of nSP is established if the patient meets (1) and satisfies any of the symptoms in (2) and (3).

The patient was diagnosed with SP, which includes the major diagnostic criteria: (1) Respiratory failure requiring mechanical ventilation; (2) Septic shock with need for vasopressors; Additionally, six minor criteria also contribute to the diagnostic framework: (1) Respiratory rate ≥30 breaths per minute; (2) PaO_2/FiO_2 ratio ≤250 mm Hg; (3) Multilobar infiltrates in the lungs; (4) Confusion or disorientation; (5) Blood urea nitrogen ≥20 mg/dL; (6) Hypotension requiring aggressive fluid resuscitation. A diagnosis of SP is established if the patient meets one major criterion or three or more minor criteria.

2.4. Data collection and definition of variables

Initially, a total of 221 pneumonia patients were included in the study. However, 17 samples were excluded due to hemolysis, lipemia, or spillage. Ultimately, 204 patients were included in model construction, comprising 100 SP patients and 104 nSP patients. Fig. 1 illustrates the screening process. Patient-related demographic information such as sex and age were recorded. The following biomarker was collected: (1) immune cells: counts of WBC, lymphocyte, T-lymphocyte, B-lymphocyte, and NK-lymphocyte; percentage of monocyte, $CD14^+HLA\text{-}DR^-$ monocyte, $CD4^+$ T-lymphocyte, $CD8^+$ T-lymphocyte, neutrophil, and lymphocyte. (2) inflammatory indicators: CRP and nCD64 index. (3) Cytokines: IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p70, IL-17, IFN-γ, IFN-α. Additionally, we defined several variables, such as the $CD4^+/CD8^+$ ratio, NLR, nCD64 index, IL-6/IL-10 ratio, and IL-4/IFN-γ ratio, which were calculated by the following formulas:

$$CD4^+/CD8^+ \text{ ratio} = \frac{\text{percentage of } CD4^+ \text{ T-lymphocyte}}{\text{percentage of } CD8^+ \text{ T-lymphocyte}}$$

$$\text{NLR} = \frac{\text{percentage of neutrophil}}{\text{percentage of lymphocyte}}$$

$$\text{nCD64 index} = \frac{\text{CD64MFI of neutrophil} \div \text{CD64 MFI of lymphocyte}}{\text{CD64 MFI of monocyte} \div \text{CD64 MFI of neutrophil}}$$

$$\text{IL-4} \Big/ \text{IFN-}\gamma \text{ ratio} = \frac{\text{concentration of IL-4}}{\text{concentration of IFN-}\gamma}$$

$$\text{IL-6} \Big/ \text{1L-10 ratio} = \frac{\text{concentration of IL-6}}{\text{concentration of IL-10}}$$

2.5. Predictors screening

To efficiently select the relevant predictors, we optimized the data and selected the most appropriate variables as predictors. First, all variables were divided into two groups according to whether the patient had SP, followed by a comparison of demographic and laboratory characteristics between these groups (Table 1). Variables demonstrating a statistically significant difference, denoted by a *P*-value of <0.05, were chosen for incorporation into the least absolute shrinkage and selection operator (LASSO) algorithm. This

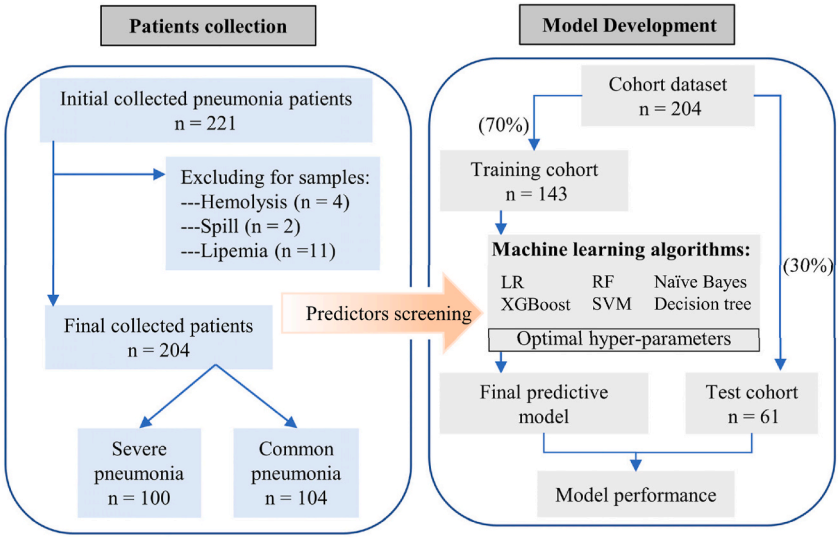


Fig. 1. Flowchart of the study. The work can be divided into two parts: Patient collection and Model development.

Table 1
Patient demographics and clinical laboratory examination characteristics of all patients.

Variables	Overall n = 204	Severe pneumonia n = 100	Non-severe pneumonia n = 104	*P* value
Age (year)	69 (53, 82)	79 (64, 84)	59 (50, 75)	<0.001
Male (n, %)	106 (51.96)	64 (64.00)	42 (40.38)	0.001
Immune cells				
WBC ($\times 10^9$/L)	7.85 (5.54, 10.54)	9.92 (7.18, 12.72)	6.21 (5.01, 8.05)	<0.001
Lymphocyte (cells/μL)	1130 (813, 1600)	950 (745, 1450)	1285 (925, 1675)	0.001
T-lymphocyte (cells/μL)	796 (572, 1104)	721 (510, 1064)	894 (666, 1129)	0.006
B-lymphocyte (cells/μL)	85 (44, 147)	73 (36, 128)	103 (65, 173)	0.001
NK-lymphocyte (cells/μL)	175 (96, 278)	155 (96, 268)	185 (97, 303)	0.336
Monocyte (%)	11.4 (8.34, 15.42)	10.74 (7.92, 14.75)	11.71 (8.39, 15.52)	0.369
$CD14^+HLA\text{-}DR^-$ monocyte (%)	0.19 (0.10, 0.49)	0.21 (0.11, 0.51)	0.18 (0.09, 0.44)	0.229
$CD4^+/CD8^+$ ratio	1.48 (1.02, 2.15)	1.30 (0.89, 1.95)	1.58 (1.15, 2.28)	0.025
NLR	5.04 (2.69, 7.85)	6.94 (4.55, 11.46)	3.35 (2.26, 5.18)	<0.001
Inflammatory indicators				
nCD64 index	0.79 (0.43, 174)	1.00 (0.54, 2.13)	0.68 (0.36, 1.58)	0.004
CRP (mg/L)	36.76 (7.08, 86.86)	66.33 (33.28, 125.10)	11.10 (2.27, 48.52)	<0.001
Serum cytokines				
IL-1β (pg/mL)	6.56 (3.79, 12.44)	7.55 (4.12, 15.09)	6.20 (3.45, 9.89)	0.019
IL-2 (pg/mL)	4.77 (2.60, 8.89)	5.06 (3.05, 10.53)	4.73 (2.36, 7.81)	0.155
IL-4 (pg/mL)	6.12 (3.48, 9.59)	6.31 (3.51, 10.29)	5.85 (3.51, 9.24)	0.377
IL-5 (pg/mL)	7.11 (4.33, 12.11)	7.77 (5.08, 12.23)	6.15 (3.89, 11.96)	0.083
IL-6 (pg/mL)	34.79 (15.66, 127.75)	110.68 (34.79, 265.28)	18.12 (11.05, 38.79)	<0.001
IL-8 (pg/mL)	78.17(40.90, 174.59)	100.31(54.05, 224.41)	63.91(37.88, 117.87)	0.001
IL-10 (pg/mL)	13.49 (7.90, 26.85)	15.59 (11.10, 30.43)	9.89 (6.46, 20.69)	0.001
IL-12p70 (pg/mL)	5.32 (3.29, 5.32)	6.43 (3.85, 9.42)	4.58 (2.92, 7.66)	0.026
IL-17 (pg/mL)	8.16(5.27, 13.07)	8.34(5.94, 13.63)	7.52(4.26, 13.07)	0.123
IFN-γ (pg/mL)	5.89 (3.74, 12.97)	7.05 (4.51, 15.46)	5.43 (3.05, 10.40)	0.012
IFN-α(pg/mL)	2.36 (1.30, 3.75)	2.62 (1.49, 4.08)	2.12 (1.03, 3.44)	0.042
TNF-α (pg/mL)	5.32 (3.33, 9.89)	6.90 (3.66, 10.72)	4.61 (2.93, 7.77)	0.002
IL-4/IFN-γ ratio	1.01 (0.46, 1.50)	0.99 (0.39, 1.38)	1.12 (0.53, 1.61)	0.153
IL-6/IL-10 ratio	2.98(1.43, 8.12)	6.06(2.59, 15.03)	1.67(0.90, 3.19)	<0.001

Abbreviation: WBC, white blood cell; NLR, neutrophile-to-lymphocyte ratio; $CD4^+/CD8^+$, $CD4^+$ T-lymphocyte percentage to $CD8^+$ T-lymphocyte percentage ratio; CRP, C-reaction protein; nCD64, neutrophile CD64; IL, interleukin; IL-4/IFN-γ, interleukin 4 to interferon γ ratio; IL-6/IL-10, interleukin 6 to interleukin 10 ratio.

algorithm, utilizing an optimal lambda parameter, effectively screened the predictors. Finally, we established the predictive model using the final predictors selected by LASSO regression in conjunction with clinical expertise. In this study, we ensured that the variables used in the models adhered to the assumption of independence by evaluating multicollinearity among predictors. Multicollinearity was assessed using the variance inflation factor (VIF) method. A VIF value greater than 5 indicates significant multicollinearity, which can violate the independence assumption.

2.6. Model development and evaluation

All enrolled patients were randomly divided into two cohorts, 70 % of patients constituted the training cohort (N = 143) for hyperparameter optimization, while the remaining 30 % formed the test cohort (N = 61) for model validation. Six machine learning algorithms, including Logistic Regression (LR), Random Forest (RF), Naïve Bayes (NB), eXtreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), and Decision Tree (DT), were employed to develop the prediction model. Each model underwent optimization through ten-fold cross-validation, tuning to the optimal parameters for final model training. The model development process is depicted in Fig. 1. During model training, the cut-off value for predicting the occurrence of SP was identified according to the Youden's-index maximum principle. A confusion matrix is a cross-table that describes the prediction results of each model by the cut-off value. Based on the metrics of the confusion matrix, parameters such as accuracy, sensitivity, and specificity were calculated to assess model performance using the following calculation formula [19]. Model evaluation in both the training and test cohorts was conducted using the area under the curve (AUC) of the ROC curve, accuracy, sensitivity, specificity, and F1_score (higher is better). Besides, the calibration curve and DCA were used to evaluate the fit and applicability of the optimal model. In addition to the model evaluation, the importance of selected predictors was ranked using SHapley Additive exPlanation (SHAP) values within the optimal prediction model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{F1\_score} = \frac{2 \times \text{precision} \times \text{sensitivity}}{\text{precision} + \text{sensitivity}}$$

2.7. Statistical analysis

Categorical variables were presented as numbers and percentages and analyzed using the Chi-square test or Fisher's exact test. Continuous variables with a normal distribution were expressed as mean ± standard deviation, and differences between them were analyzed using the two-tailed Student's t-tests, while the non-normally distributed variables were expressed as median with interquartile range (IQR) and analyzed using the Mann-Whitney *U* test. Statistical significance was defined as *P* values < 0.05 for all analyses. Statistical analysis of this study was conducted using R software (version 3.6.1).

3. Results

3.1. Patient characteristics and variables

A total of 204 patients were enrolled in this study, of whom 100 were diagnosed as SP. As shown in Table 1, SP patients exhibited a higher prevalence of elderly individuals and males. Most variables of immune cells, inflammatory indicators, and serum cytokines were also significantly elevated in SP patients compared to nSP. These included WBC count, NLR, nCD64 index, CRP, IL-1β, IL-6, IL-8, IL-10, IL-12p7, IFN-γ, IFN-α, TNF-α, IL-6/IL-10 ratio, while lymphocyte count, T-lymphocyte count, B-lymphocyte count, and $CD4^+/CD8^+$ ratio were lower. Nevertheless, variables such as NK-lymphocyte count, monocyte percentage, $CD14^+HLA\text{-}DR^-$ monocyte percentage, IL-2, IL-4, IL-5, IL-17, and IL-4/IFN-γ ratio did not show significant differences between the two groups ($P > 0.05$). Ultimately, variables including age, sex, WBC count, T-lymphocyte count, NLR, CRP, TNF-α, IL-4/IFN-γ ratio, and IL-6/IL-10 ratio were selected as potential predictors using LASSO regression and clinical experience. A heatmap of correlations revealed no significant correlations among the nine variables (Fig. 2). Multicollinearity was tested among the final set of predictors using the VIF method. The VIF values for the selected variables were all below 5, indicating the absence of multicollinearity (Table S1). The experimental detection time for these variables is 2–3 h, which is sufficient to meet the clinical monitoring needs for severe pneumonia (Table S2). Subsequently, all patients were randomly divided into a training cohort and a test cohort at a 7:3 ratio. Statistical analysis indicated

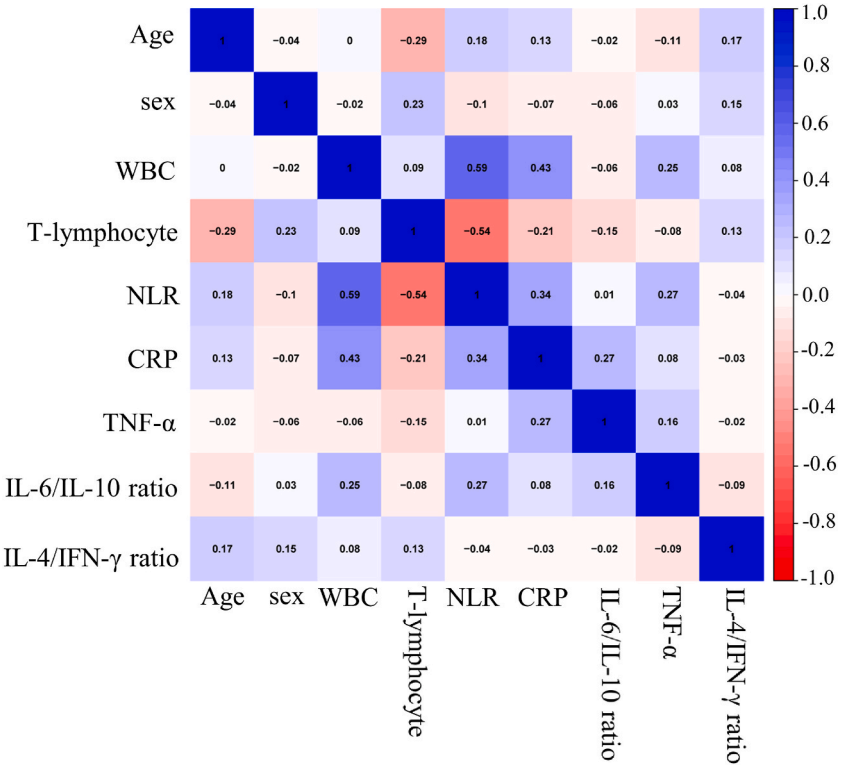


Fig. 2. Heatmap of correlations among the selected variables.

that most factors were similar between the two cohorts ($P > 0.05$), ensuring comparability for further analysis (Table S3).

3.2. Parameter optimization and model development

We established the predictive models using the nine potential variables with six machine learning algorithms as mentioned earlier. The specific parameters optimization for all models are detailed in Table S4. In the training cohort, the confusion matrix and ROC curve for each model were shown in Fig. S1 and Fig. S2, respectively. The model performance metrics, as shown in Table 2, including accuracy (ranging from 0.776 to 0.979), sensitivity (ranging from 0.765 to 0.985), specificity (ranging from 0.707 to 0.973), F1_Score (ranging from 0.784 to 0.978), and AUC (ranging from 0.862 to 0.996) with 95 % confidence intervals (*CI*). These results indicate excellent performance across all models. However, model performance evaluation was focused on comparing those metrics in the test cohort. Confusion matrices for the test cohort are shown in Fig. 3. The XGBoost model achieved the best diagnostic performance, with an AUC of 0.901 (95 % *CI*: 0.827–0.985), outperformed LR, RF, SVM, NB, and DT models with AUCs of 0.808 (95 % *CI*: 0.701–0.916), 0.891 (95 % *CI*: 0.810–0.972), 0.786 (95 % *CI*: 0.663–0.908), 0.763 (95 % *CI*: 0.644–0.882) and 0.659 (95 % *CI*: 0.527–0.792), respectively (Fig. 4). Detailed metrics of all models were summarized in Table 3. Among them, the XGBoost model also exhibited the highest accuracy (0.803), sensitivity (0.844), and F1_score (0.818), albeit with lower specificity (0.759) compared to the DT model (0.862) and the RF model (0.862). The basic structure and algorithm steps of XGBoost are shown in Fig. S3 and Table S5, respectively.

Based on the aforementioned results, we selected the XGBoost model as the final model due to its superior overall performance in the test cohort. In Fig. 5A, the calibration curve of the XGBoost model is closely aligned with the reference line, indicating good predictive consistency between predicted probability and actual probability. Subsequently, DCA was used to assess the clinical net benefit of the XGBoost model in the context of decision-making. As shown in Fig. 5B, the result shows higher overall net benefits within a range of threshold probabilities between 14 % and 100 %, suggesting that the XGBoost model might offer greater utility in guiding clinical decisions across a wide range of scenarios.

3.3. Individual variable importance

To explain the importance of each selected predictor in the optimal model, the SHAP algorithm was performed to rank the predictors based on their influence on the model outcome (Fig. 6). Therefore, we can observe that the feature importance ranking in the XGBoost model is as follows: WBC, Age, CRP, IL-6-IL-10 ratio, T-lymphocyte count, sex, IL-4/IFN-γ ratio, TNF-α, and NLR. Notably, elevated levels of WBC count, age, CRP, NLR, and the IL-6/IL-10 ratio, along with decreased levels of T-lymphocyte count and the IL-4/IFN-γ ratio, were associated with a higher risk of SP. Additionally, being male was also found to be a factor in increasing the risk of SP.

4. Discussion

Accumulating evidence underscores a strong association between inflammation and the risk of respiratory diseases [20]. Inflammation-associated factors have emerged as promising biomarkers for predicting the development of pneumonia, based on both quantitative and qualitative evaluations [21]. Recently, novel inflammation biomarkers, such as cytokines and various cell subsets in peripheral blood, have garnered attention for their potential predictive capability in infectious diseases [22]. However, the predictive value of these variables for SP patients remains to be fully elucidated. In our study, we developed six ML prediction models for SP, using the variables related to immune cells, inflammatory indicators, serum cytokines, and clinical characteristics. Following the predictors screening, age, sex, WBC count, T-lymphocyte count, NLR, CRP, TNF-α, IL-4/IFN-γ ratio, and IL-6/IL-10 ratio were included. Importantly, the XGBoost model emerges as the superior choice, outperforming LR, RF, SVM, NB, and DT models. As a result, the XGBoost model holds significant potential to aid clinical decision-making for pneumonia patients and improve SP prediction in clinical practice.

Among the predictors screened in this study, we initially analyzed all variables between the SP and nSP groups. Our findings revealed that advanced age was associated with a higher susceptibility to developing SP, which is consistent with previous studies [23–25]. This association may be attributed to age-related immune weakening in elderly patients, where declining immune function can increase their vulnerability to infections. Furthermore, we observed a higher proportion of male patients in the SP group, which we speculate may be linked to smoking habits among males. Regarding laboratory-associated, the increased levels of WBC count, NLR, and CRP were common features among severe pneumonia patients compared to those with milder pneumonia [5,7]. Our findings were

Table 2
Summary of specific metrics of the six machine algorithm models in the training cohort.

Model	Accuracy	Sensitivity	Specificity	F1_Score	AUC (95%*CI*)
LR	0.860	0.838	0.880	0.851	0.920 (0.877–0.962)
RF	0.902	0.882	0.920	0.896	0.952 (0.920–0.984)
SVM	0.846	0.765	0.920	0.825	0.896 (0.843–0.950)
XGBoost	0.979	0.985	0.973	0.978	0.996 (0.992–1.000)
NB	0.776	0.853	0.707	0.784	0.862 (0.804–0.920)
DT	0.888	0.838	0.933	0.877	0.936 (0.896–0.976)

Abbreviation: LR, Logistic Regression; RF, Random Forest; NB, Naïve Bayes; XGBoost, eXtreme Gradient Boosting; SVM, Support Vector Machine; DT, Decision Tree; AUC, area under the receiver operating characteristic curve; 95 % *CI*, 95 % confidence interval.

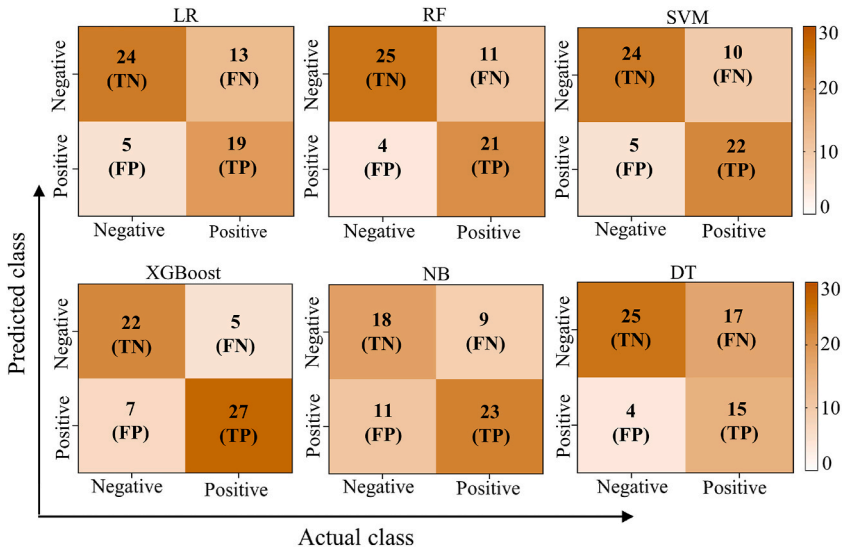


Fig. 3. Confusion matrix of all predictive models in the test cohort. The test sample was oversampled from the original cohort dataset (30% of the cohort). The horizontal axis indicated the true label of positive and negative for severe pneumonia (TN and TP mapped classes where each model predicted the class correctly as negative and positive severe pneumonia, respectively). The vertical axis indicated the predictive label for each algorithm (FN and FP mapped classes where each model predicted the class incorrectly as negative and positive severe pneumonia, respectively).

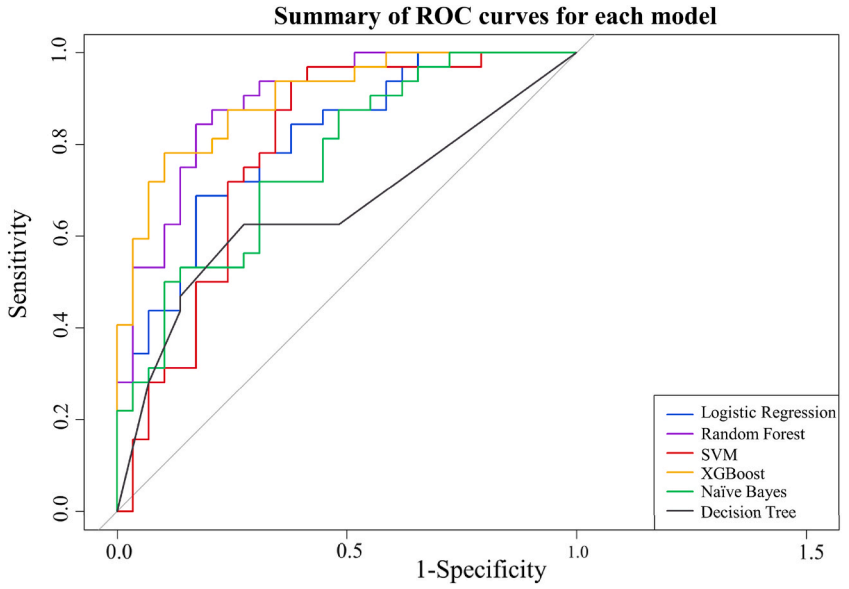


Fig. 4. Test cohort ROC curve for each model.

Table 3
Summary of specific metrics of the six machine algorithm models in the test cohort.

Model	Accuracy	Sensitivity	Specificity	F1 Score	AUC (95%*CI*)
LR	0.705	0.594	0.828	0.679	0.808 (0.701–0.916)
RF	0.754	0.656	0.862	0.737	0.891 (0.810–0.972)
SVM	0.754	0.688	0.828	0.746	0.786 (0.663–0.908)
XGBoost	0.803	0.844	0.759	0.818	0.901 (0.827–0.975)
NB	0.672	0.719	0.621	0.697	0.763 (0.644–0.882)
DT	0.656	0.469	0.862	0.588	0.659 (0.527–0.792)

Abbreviation: LR, Logistic Regression; RF, Random Forest; NB, Naïve Bayes; XGBoost, eXtreme Gradient Boosting; SVM, Support Vector Machine; DT, Decision Tree; AUC, area under the receiver operating characteristic curve; 95 % *CI*, 95 % confidence interval.

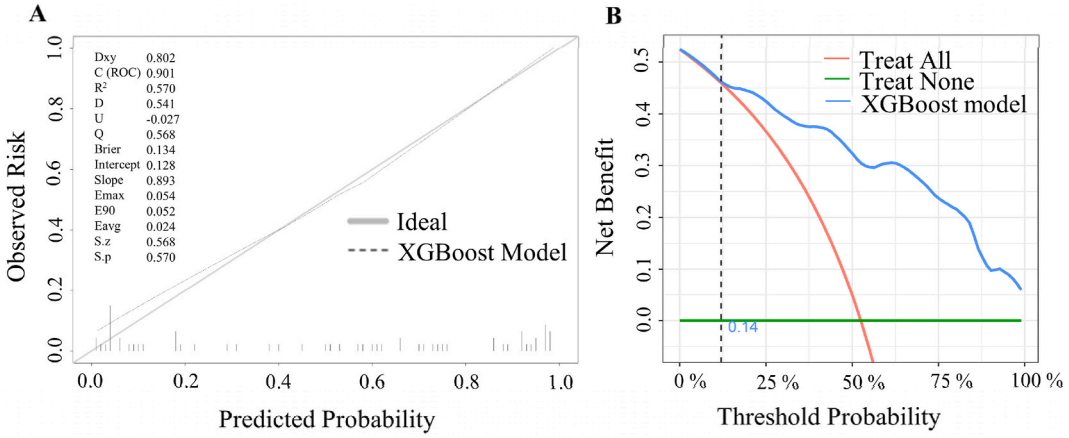


Fig. 5. Calibration curves and DCA of XGBoost model in the testing cohort. A calibration curve, B DCA, decision curve analysis.

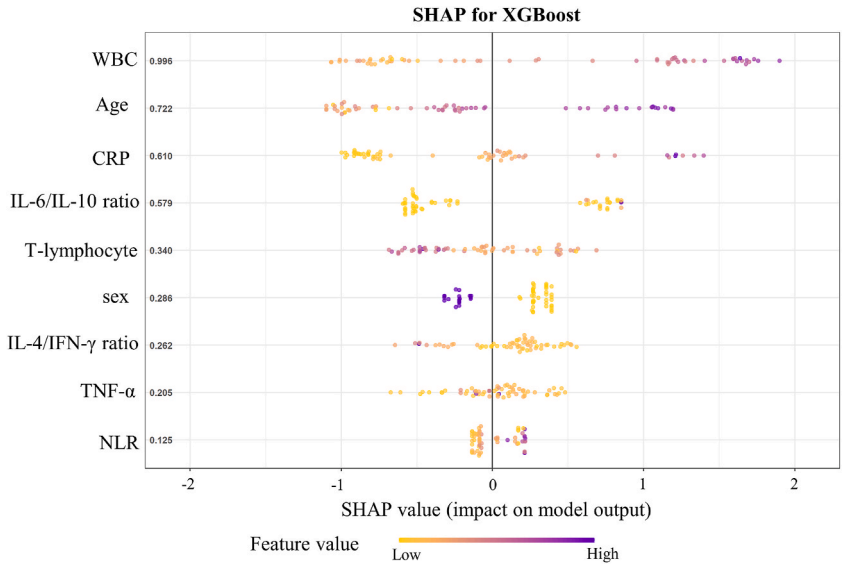


Fig. 6. SHAP summary plot of the nine variables in the XGBoost model. SHAP, Shapley additive explanation.

consistent with the above-reported results. Notably, our findings showed a significant elevation of the nCD64 index in the severe pneumonia group. Burgo et al. also found that the increased nCD64 was a risk factor for ICU admission and clinical deterioration in CAP patients, although its accuracy in predicting poor prognosis was not high [26]. Thus, we suggest that the nCD64 index may serve as a valuable supplementary detection marker in suspected pneumonia cases, despite it being considered as one of the specific indicators for sepsis [27]. Besides, our data also revealed lymphopenia as an important clinical feature in SP patients, characterized by the low-count level of total lymphocyte, T-lymphocyte, and B-lymphocyte, among which the T-lymphocyte count was identified as a predictor included in the final model. In contrast, the NK-lymphocyte count did not yield similar results, suggesting a potentially more nuanced role for different lymphocyte subsets in the pathogenesis of SP. Torres et al. [28,29] have identified lymphopenia (total lymphocyte count <724 cells/μL) in community-acquired pneumonia (CAP) patients as a significant risk factor for mortality. Their study highlights a strong correlation between decreased total lymphocyte counts and specifically reduced T-lymphocyte populations, while associations with NK-lymphocytes and B-lymphocytes were more modest. Therefore, we hypothesize that T-lymphocytopenia is the primary driver of lymphocytopenia observed in severe CAP cases. Moreover, our findings reveal a marked decrease in T-lymphocyte count, coupled with a substantial decline in the $CD4^+/CD8^+$ ratio among SP patients. These observations underscore the crucial significance of detecting immune cell subsets in individuals predisposed to developing SP, highlighting the necessity of such assessments for a better understanding of disease progression and risk assessment.

Serum cytokines, significant soluble proteins secreted by immune cells and nonimmune cells, play a crucial role in regulating cell growth, differentiation, and effects through receptor binding, thereby modulating immune responses. In the context of inflammation, cytokines function as regulators, maintaining a delicate balance between pro-inflammatory and anti-inflammatory processes. Our

findings revealed that SP patients exhibited elevated levels of IL-1β, IL-6, IL-8, IL-10, IL-12p70, IFN-γ, and IFN-α, indicating a heightened inflammatory status compared to nSP patients. Recent studies on COVID-19 have highlighted IL-6 as a core pro-inflammatory cytokine contributing to the phenomenon of "cytokine storm" [10,30]. Conversely, IL-10, an anti-inflammatory factor, tends to increase significantly in the later stages of the disease, inhibiting the over-secretion of pro-inflammatory factors [31,32]. Furthermore, we observed a significant increase in the IL-6/IL-10 ratio among SP patients, suggesting that this ratio may serve as an indicator of the progression from common pneumonia to SP. This underscores the significance of assessing inflammation status as an early biomarker for pneumonia severity. The IL-4/IFN-γ ratio is recognized as an expression product of $CD4^+$ T helper (Th) cells, reflecting the balanced immune response between Th1 and Th2 subsets [33]. While neither the IL-6/IL-10 nor the IL-4/IFN-γ ratio was directly selected by LASSO regression for model construction, evaluating inflammatory profiles remains vital. Consequently, our predictive models were established using a comprehensive set of predictors, including age, sex, WBC count, T-lymphocyte count, NLR, CRP, TNF-α, the IL-6/IL-10 ratio, and the IL-4/IFN-γ ratio.

In recent years, ML algorithms have been widely used in developing disease prediction models due to their ability to handle complex and high-dimensional data, thereby improving the accuracy of these models [34,35]. We performed six ML algorithms to build prediction models, which included a linear model (LR), non-linear models (RF, XGBoost, and SVM), Naïve Bayes, and Decision Tree. Our study suggested that the XGBoost model demonstrated the best diagnostic performance for SP, with an AUC of 0.901 (95 % *CI*: 0.827–0.985), outperformed LR, RF, SVM, NB, and DT models. Additionally, the XGBoost model also exhibited the highest accuracy (0.803), sensitivity (0.844), and F1_score (0.818), although it had lower specificity (0.759) compared to the DT model (0.862) and the RF model (0.862). XGBoost, a supervised machine learning algorithm known for enhancing prediction accuracy and efficiency by amalgamating diverse tree learners, has been a robust method with good performance when dealing with complex relationships and nonlinear data [36,37]. Due to its resistance to overfitting, XGBoost is particularly effective for datasets with imbalanced feature/outcome ratios and offers more flexibility in tuning hyperparameters compared to other algorithms [38]. Hong et al. used two ML algorithms, XGBoost and RF, to develop a prediction model for severe COVID-19 pneumonia based on cytokines and immune cell profile, and recommended XGBoost as the best model due to its highest discriminatory performance [39]. An XGBoost-based model that incorporates key hematological indicators like CRP, PCT, albumin, hemoglobin, and progressive dyspnea, which outperforms RF, LR, and SVM models in predicting severe *pneumocystis carinii* pneumonia by demonstrating higher specificity [40]. To elucidate the impact value of all selected predictors in the XGBoost model, we performed a SHAP analysis. The SHAP plot revealed the top three predictors in the XGboost mode were WBC count, age, and CRP. These top predictive features all have clinically plausible explanations.

Nevertheless, this study has several limitations. Firstly, the sample size was relatively limited, and it is necessary to validate the SP prediction model in a larger population. Secondly, the diagnostic performance of the prediction model for severe pneumonia needs to be assessed by comparing it with established evaluation systems such as CURB and PSI scores. Finally, certain clinical features and laboratory-associated indicators, such as respiratory symptoms, and pathogen type of infection, were not included in this study, potentially limiting the validation of our model. Based on the aforementioned results, this study constructs an ML-based prediction model for anticipating severe progression in pneumonia patients using a more comprehensive array of inflammatory biomarkers collected from multiple centers.

5. Conclusions

In summary, based on XGBoost, we constructed a practical prediction model for severe pneumonia based on inflammatory biomarkers and demonstrated its good effectiveness through ROC analysis, calibration curve analysis, and DCA. Furthermore, this SP prediction model has an important clinical potential for assessing the risk of SP in patients with pneumonia, thereby facilitating the implementation of early prevention strategies. In the future, we will design a prospective study to validate the proposed model on a larger dataset while simultaneously collecting and organizing more comprehensive patient information to establish multiple models.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics statement

This study was conducted with the approval of the Ethics Committee of Chongqing University Cancer Hospital (Ethical code: CZLS2022022-A-19) and met the Helsinki Declaration. All participants/patients (or their proxies/legal guardians) provided informed consent to participate in the study.

Funding

This work was supported by the Science-Health Joint Medical Scientific Research Project of Chongqing (Grant number. 2021MSXM272).

Data availability statement

The raw data underlying this paper are available upon request to the corresponding author.

CRediT authorship contribution statement

Zailin Yang: Writing – review & editing, Writing – original draft, Investigation, Funding acquisition. **Shuang Chen:** Writing – review & editing, Writing – original draft, Software, Methodology. **Xinyi Tang:** Software, Data curation. **Jiao Wang:** Data curation. **Ling Liu:** Data curation. **Weibo Hu:** Data curation. **Yulin Huang:** Data curation. **Jian'e Hu:** Data curation. **Xiangju Xing:** Data curation. **Yakun Zhang:** Data curation. **Jun Li:** Data curation. **Haike Lei:** Visualization, Validation, Supervision, Software, Methodology, Conceptualization. **Yao Liu:** Visualization, Supervision, Investigation, Conceptualization.

Declaration of competing interest

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

Acknowledgment

We thank the technical support and statistical analysis support from Mingjie Wang. He contributed to the advancement of our work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e37367.

References

[1] A. Torres, C. Cilloniz, M.S. Niederman, R. Menéndez, J.D. Chalmers, R.G. Wunderink, et al., Pneumonia, Nat. Rev. Dis. Prim. 7 (2021) 25, https://doi.org/10.1038/s41572-021-00259-0.
[2] F. Ariani, K. Liu, Z. Jing, J. Qu, Glucocorticosteroid in treatment of severe pneumonia, Mediat. Inflamm. 2013 (2013) 865635, https://doi.org/10.1155/2013/865635.
[3] S.B. Davide Leoni, Evdoxia Tsigou, Despoina Koulenti, What we learned from the EU-VAP/CAP study for severe pneumonia, Clin. Pulm. Med. 24 (2017) 112–120, https://doi.org/10.1097/CPM.0000000000000204.
[4] C. Marti, N. Garin, O. Grosgurin, A. Poncet, C. Combescure, S. Carballo, et al., Prediction of severe community-acquired pneumonia: a systematic review and meta-analysis, Crit. Care 16 (2012) R141, https://doi.org/10.1186/cc11447.
[5] E. Cataudella, C.M. Giraffa, S. Di Marca, A. Pulvirenti, S. Alaimo, M. Pisano, et al., Neutrophil-to-lymphocyte ratio: an emerging marker predicting prognosis in elderly adults with community-acquired pneumonia, J. Am. Geriatr. Soc. 65 (2017) 1796–1801, https://doi.org/10.1111/jgs.14894.
[6] Â.G. Vasconcellos, J. Clarêncio, D. Andrade, M.-R.A. Cardoso, A. Barral, C.M. Nascimento-Carvalho, Systemic cytokines and chemokines on admission of children hospitalized with community-acquired pneumonia, Cytokine 107 (2018) 1–8, https://doi.org/10.1016/j.cyto.2017.11.005.
[7] Z. Li, N. Jiang, X. Li, B. Yang, M. Jin, Y. Sun, et al., Two novel nomograms based on inflammatory cytokines or lymphocyte subsets to differentially diagnose severe or critical and Non-Severe COVID-19, Aging 13 (2021) 17961–17977, https://doi.org/10.18632/aging.203307.
[8] Y. Zhang, S. Mei, Y. Zhou, M. Huang, G. Dong, Z. Chen, Cytokines as the good predictors of refractory Mycoplasma pneumoniae pneumonia in school-aged children, Sci. Rep. 6 (2016) 37037, https://doi.org/10.1038/srep37037.
[9] J. Haugen, R.K. Chandyo, K.A. Brokstad, M. Mathisen, M. Ulak, S. Basnet, et al., Cytokine concentrations in plasma from children with severe and non-severe community acquired pneumonia, PLoS One 10 (2015) e0138978, https://doi.org/10.1371/journal.pone.0138978.
[10] H. Han, Q. Ma, C. Li, R. Liu, L. Zhao, W. Wang, et al., Profiling serum cytokines in COVID-19 patients reveals IL-6 and IL-10 are disease severity predictors, Emerg. Microb. Infect. 9 (2020) 1123–1130, https://doi.org/10.1080/22221751.2020.1770129.
[11] L.D. Chen, Z.Y. Zhang, X.J. Wei, Y.Q. Cai, W.Z. Yao, M.H. Wang, et al., Association between cytokine profiles and lung injury in COVID-19 pneumonia, Respir. Res. 21 (2020) 201, https://doi.org/10.1186/s12931-020-01465-2.
[12] M.S. Paats, I.M. Bergen, W.E.J.J. Hanselaar, E.C. Groeninx van Zoelen, H.C. Hoogsteden, R.W. Hendriks, et al., Local and systemic cytokine profiles in nonsevere and severe community-acquired pneumonia, Eur. Respir. J. 41 (2013) 1378–1385, https://doi.org/10.1183/09031936.00060112.
[13] B. Lu, M. Liu, J. Wang, H. Fan, D. Yang, L. Zhang, et al., IL-17 production by tissue-resident MAIT cells is locally induced in children with pneumonia, Mucosal Immunol. 13 (2020) 824–835, https://doi.org/10.1038/s41385-020-0273-y.
[14] H.-L. Jin, L. Zhan, S.-F. Mei, Z.-Y. Shao, Serum cytokines and FeNO in school-aged children with mycoplasma pneumoniae pneumonia, Med. Sci. Mon. Int. Med. J. Exp. Clin. Res. 26 (2020) e923449, https://doi.org/10.12659/MSM.923449.
[15] Q. Liu, Y. Gao, Q. Ou, Y. Xu, Z. Zhou, T. Li, et al., Differential expression of CD64 in patients with Mycobacterium tuberculosis infection: a potential biomarker for clinical diagnosis and prognosis, J. Cell Mol. Med. 24 (2020) 13961–13972, https://doi.org/10.1111/jcmm.16004.
[16] C.E.M. Viana, D.M. Matos, MdF. Oliveira, A.C. da Costa, TPdA. Filho, P.A.M. Filho, et al., Immunosuppressive cd14+/HLA-DRlow/− monocytes in patients with Chagas disease, Acta Trop. 224 (2021) 106154, https://doi.org/10.1016/j.actatropica.2021.106154.
[17] M. Lagadinou, N. Zareifopoulos, D. Gkentzi, F. Sampsonas, E. Kostopoulou, M. Marangos, et al., Alterations in lymphocyte subsets and monocytes in patients diagnosed with SARS-CoV-2 pneumonia: a mini review of the literature, Eur. Rev. Med. Pharmacol. Sci. 25 (2021) 5057–5062, https://doi.org/10.26355/eurrev_202108_26463.
[18] J.P. Metlay, G.W. Waterer, A.C. Long, A. Anzueto, J. Brozek, K. Crothers, et al., Diagnosis and treatment of adults with community-acquired pneumonia. An official clinical practice guideline of the American thoracic society and infectious diseases society of America, Am. J. Respir. Crit. Care Med. 200 (2019) e45–e67, https://doi.org/10.1164/rccm.
[19] T. Abd El-Hafeez, M.Y. Shams, Y.A.M.M. Elshaier, H.M. Farghaly, A.E. Hassanien, Harnessing machine learning to find synergistic combinations for FDA-approved cancer drugs, Sci. Rep. 14 (2024) 2428, https://doi.org/10.1038/s41598-024-52814-w.
[20] S.Y. Jae, K.S. Heffernan, S. Kurl, S.K. Kunutsor, C.-H. Kim, B.D. Johnson, et al., Cardiorespiratory fitness, inflammation, and the incident risk of pneumonia, J Cardiopulm Rehabil Prev 41 (2021) 199–201, https://doi.org/10.1097/HCR.0000000000000581.
[21] J.P. Mizgerd, Inflammation and pneumonia: why are some more susceptible than others? Clin. Chest Med. 39 (2018) 669–676, https://doi.org/10.1016/j.ccm.2018.07.002.
[22] H. Akbari, R. Tabrizi, K.B. Lankarani, H. Aria, S. Vakili, F. Asadian, et al., The role of cytokine profile and lymphocyte subsets in the severity of coronavirus disease 2019 (COVID-19): a systematic review and meta-analysis, Life Sci. 258 (2020) 118167, https://doi.org/10.1016/j.lfs.2020.118167.

[23] Y. Song, X. Wang, K. Lang, T. Wei, J. Luo, Y. Song, et al., Development and validation of a nomogram for predicting 28-day mortality on admission in elderly patients with severe community-acquired pneumonia, J. Inflamm. Res. 15 (2022) 4149–4158, https://doi.org/10.2147/JIR.S369319.
[24] A. Bonaventura, A. Vecchié, T.S. Wang, E. Lee, P.C. Cremer, B. Carey, et al., Targeting GM-CSF in COVID-19 pneumonia: rationale and strategies, Front. Immunol. 11 (2020) 1625, https://doi.org/10.3389/fimmu.2020.01625.
[25] L. Yang, D. He, D. Huang, Z. Zhang, Z. Liang, Development and validation of nomogram for hospital mortality in immunocompromised patients with severe pneumonia in intensive care units: a single-center, retrospective cohort study, Int. J. Gen. Med. 15 (2022) 451–463, https://doi.org/10.2147/IJGM.S344544.
[26] J. Burgos, I. Los-Arcos, D. Álvarez de la Sierra, V. Falcó, A. Aguiló, I. Sánchez, et al., Determination of neutrophil CD64 expression as a prognostic biomarker in patients with community-acquired pneumonia, Eur. J. Clin. Microbiol. Infect. Dis. 35 (2016) 1411–1416, https://doi.org/10.1007/s10096-016-2678-9.
[27] C.F. Yeh, C.C. Wu, S.H. Liu, K.F. Chen, Comparison of the accuracy of neutrophil CD64, procalcitonin, and C-reactive protein for sepsis identification: a systematic review and meta-analysis, Ann. Intensive Care 9 (2019) 5, https://doi.org/10.1186/s13613-018-0479-2.
[28] R. Méndez, R. Menéndez, I. Amara-Elori, L. Feced, A. Piró, P. Ramírez, et al., Lymphopenic community-acquired pneumonia is associated with a dysregulated immune response and increased severity and mortality, J. Infect. 78 (2019) 423–431, https://doi.org/10.1016/j.jinf.2019.04.006.
[29] J.F. Bermejo-Martin, C. Cilloniz, R. Mendez, R. Almansa, A. Gabarrus, A. Ceccato, et al., Lymphopenic community acquired pneumonia (L-CAP), an immunological phenotype associated with higher risk of mortality, EBioMedicine 24 (2017) 231–236, https://doi.org/10.1016/j.ebiom.2017.09.023.
[30] E.A. Coomes, H. Haghbayan, Interleukin-6 in Covid-19: a systematic review and meta-analysis, Rev. Med. Virol. 30 (2020) 1–9, https://doi.org/10.1002/rmv.2141.
[31] L.A. González, F. Melo-González, V.P. Sebastián, O.P. Vallejos, L.P. Noguera, I.D. Suazo, et al., Characterization of the anti-inflammatory capacity of IL-10-producing neutrophils in response to infection, Front. Immunol. 12 (2021) 638917, https://doi.org/10.3389/fimmu.2021.638917.
[32] B.M. Henry, S.W. Benoit, J. Vikse, B.A. Berger, C. Pulvino, J. Hoehn, et al., The anti-inflammatory cytokine response characterized by elevated interleukin-10 is a stronger predictor of severe disease and poor outcomes than the pro-inflammatory cytokine response in coronavirus disease 2019 (COVID-19), Clin. Chem. Lab. Med. 59 (2021) 599–607, https://doi.org/10.1515/cclm-2020-1284.
[33] X. Zhu, J. Zhu, CD4 T helper cell subsets and related human immunological disorders, Int. J. Mol. Sci. 21 (2020) 8011, https://doi.org/10.3390/ijms21218011.
[34] D.A. Abdel Hady, O.M. Mabrouk, T. Abd El-Hafeez, Employing machine learning for enhanced abdominal fat prediction in cavitation post-treatment, Sci. Rep. 14 (2024) 11004, https://doi.org/10.1038/s41598-024-60387-x.
[35] J.H. Lee, G.Y. Song, J. Lee, S.R. Kang, K.M. Moon, Y.D. Choi, et al., Prediction of immunochemotherapy response for diffuse large B-cell lymphoma using artificial intelligence digital pathology, J Pathol Clin Res 10 (2024) e12370, https://doi.org/10.1002/2056-4538.12370.
[36] Y. Guo, Y. Yang, R. Li, X. Liao, Y. Li, Cadmium accumulation in tropical island paddy soils: from environment and health risk assessment to model prediction, J. Hazard Mater. 465 (2023) 133212, https://doi.org/10.1016/j.jhazmat.2023.133212.
[37] M. Noorunnahar, A.H. Chowdhury, F.A. Mila, A tree based eXtreme Gradient Boosting (XGBoost) machine learning model to forecast the annual rice production in Bangladesh, PLoS One 18 (2023) e0283452, https://doi.org/10.1371/journal.pone.0283452.
[38] A. Vaid, L. Chan, K. Chaudhary, S.K. Jaladanki, I. Paranjpe, A. Russak, et al., Predictive approaches for acute dialysis requirement and death in COVID-19, Clin. J. Am. Soc. Nephrol. 16 (2021) 1158–1168, https://doi.org/10.2215/CJN.17311120.
[39] W. Hong, X. Zhou, S. Jin, Y. Lu, J. Pan, Q. Lin, et al., A comparison of XGBoost, random forest, and nomograph for the prediction of disease severity in patients with COVID-19 pneumonia: implications of cytokine and immune cell profile, Front. Cell. Infect. Microbiol. 12 (2022) 819267, https://doi.org/10.3389/fcimb.2022.819267.
[40] Y. Liu, T. Qiu, H. Hu, C. Kong, Y. Zhang, T. Wang, et al., Machine learning models for prediction of severe pneumocystis carinii pneumonia after kidney transplantation: a single-center retrospective study, Diagnostics 13 (2023) 2735, https://doi.org/10.3390/diagnostics13172735.