International Journal of General Medicine Dovepress

Open Access Full Text Article ORIGINAL RESEARCH

Early Recognition of Secondary Asthma Caused by Lower Respiratory Tract Infection in Children Based on Multi-Omics Signature: A Retrospective Cohort Study

Zhihui Rao, Shuqin Zhang, Wenlin Xu, Pan Huang, Xiaofei Xiao, Xiuxiu Hu

Department of Pediatric Comprehensive Internal Medicine, Jiangxi Maternal and Child Health Hospital, Nanchang, 330008, People's Republic of China

Correspondence: Zhihui Rao, Email 13870462660@163.com

Objective: To explore the types of pathogens causing lower respiratory tract infections (LTRIs) in children and construction of a predictive model for monitoring secondary asthma caused by LTRIs.

Methods: Seven hundred and seventy-five children with LTRIs treated from June 2017 to July 2024 were selected as research subjects. Bacterial isolation and culture were performed on all children, and drug sensitivity tests were conducted on the isolated pathogens; And according to whether the child developed secondary asthma during treatment, they were divided into asthma group (n = 116) and non-asthma group (n = 659); Using logistic regression model to analyze the risk factors affecting secondary asthma in children with LTRIs, and establishing machine learning (ie nomogram and decision tree) prediction models; Using ROC curve analysis machine learning algorithms to predict AUC values, sensitivity, and specificity of secondary asthma in children with LTRIs.

Results: 792 pathogenic bacteria were isolated from 775 children with LTRIs through bacterial culture, including 261 Gram positive bacteria (32.95%) and 531 Gram negative bacteria (67.05%). Logistic regression model analysis showed that Glycerophospholipids, Sphingolipids and radiomics characteristics were risk factors for secondary asthma in children with LTRIs ($P < 0.05$). The AUC, sensitivity, and specificity of nomogram prediction for secondary asthma in children with LTRIs were 0.817(95CI: 0.760–0.874), 82.3%, and 76.6%, respectively; The AUC of decision tree prediction for secondary asthma in children with LTRIs is 0.926(95% CI: 0.869–0.983), with a sensitivity of 96.7% and a specificity of 87.8%.

Conclusion: LTRIs in children are mainly caused by Staphylococcus aureus, Streptococcus pneumoniae, Staphylococcus epidermidis, Escherichia coli, Klebsiella pneumoniae, and Pseudomonas aeruginosa; In addition, machine learning combined with multi-omics prediction models has shown good ability in predicting LTRIs combined with asthma, providing a non-invasive and effective method for clinical decision-making.

Keywords: children, lower respiratory tract infection, pathogenic bacteria, radiomics, untargeted metabolomics, asthma, prediction model

Introduction

Lower respiratory tract infections (LRTIs), as the most common respiratory disease among school-age children, account for up to 4 million deaths worldwide each year.[1–3] On the one hand, the reason is that the respiratory system in children is not yet mature, such as tracheal stenosis, bronchial stenosis, poor ciliary movement, and a small number of alveoli, which can easily lead to adhesions and blockages.[4,5] On the other hand, due to the weaker resistance of children's bodies, favorable conditions are provided for the invasion of pathogens.[6] Given this, it is extremely urgent to implement effective treatment for LRTIs as soon as possible, especially in order to control the progression of the child's condition and improve prognosis. It is alarming that currently, antibiotics are used in clinical practice to treat LRTIs in children.[7] Although they can stabilize the condition, blindly using antibiotics may lead to drug resistance, thereby increasing the

Received: 2 October 2024
Accepted: 10 December 2024
Published: 14 December 2024

difficulty of treatment.[8,9] Moreover, it is highly possible to successfully induce asthma under the influence of LRTIs. Unfortunately, the pathogenesis of LRTIs combined with asthma in children is not yet clear in clinical practice, and there is a lack of effective predictive tools.

Previous scholars have attempted to explore the risk factors for asthma associated with LRTIs, focusing on factors such as the child's personal history and family history or leaning towards the inflammatory cytokine storm theory.[10–13] In fact, the mechanism of LRTIs combined with asthma is very complex, and there are significant differences among individual patients. Relying solely on medical history collection and serological clinical indicators cannot construct an accurate and reliable prediction model for LRTIs combined with asthma. Given this, it is extremely urgent to seek a new and reliable biomarker that can be used for the development of LRTIs combined with asthma prediction models. In recent years, multi-omics has made rapid progress in medical clinical practice. For example, metabolomics, as an important component of systems biology, can be used to identify and quantify metabolites in biological samples or analyze their metabolic pathways.[14,15] Radiomics, as an emerging field of quantitative image analysis, aims to correlate large-scale extracted image information with clinical and biological endpoints.[16] However, the mechanism exploration and predictive biomarker exploration of LRTIs combined with asthma in pediatric patients based on metabolomics and radiomics have not yet been implemented. Previous studies have shown that metabolomics is highly suitable for addressing this limitation through comprehensive analysis of metabolites and plays a positive role in biological signal transduction related to respiratory diseases.[17] Encouraged by this, we hope to construct a multimodal LRTIs combined with asthma candidate biomarker mining through multi-omics (ie metabolomics, radiomics), in order to construct highly accurate personalized LRTIs combined with asthma clinical prediction model.

In this study, we retrospectively analyzed candidate biomarkers for plasma metabolomics and non-enhanced CT imaging differences in children with LRTIs and used machine learning algorithms to construct multimodal LRTIs combined with asthma early prediction model, in order to provide more clues for early diagnosis and health monitoring of LRTIs combined with asthma, as well as experimental evidence for elucidating metabolic differences in disease progression and clinical decision-making.

Materials and Methods

Study Population

We retrospectively selected 775 children with LRTIs who were diagnosed and treated from June 2017 to July 2024 as the research subjects. Bacterial isolation and culture were performed on all children, and drug sensitivity tests were conducted on the isolated pathogens. According to whether the child has secondary asthma during treatment, they are divided into asthma group and non-asthma group. Inclusion criteria: (1) Children diagnosed with LRTIs clinically and exhibiting clinical symptoms such as pulmonary rales, fever, and cough; (2) Children aged 3–7 years old; (3) Children with first-time onset of illness; (4) Children with clinical data and follow-up cooperation. Exclusion criteria: (1) Children with important organ dysfunction; (2) Children with congenital heart disease; (3) Children who have already been treated with antibiotics before admission. The diagnostic criteria for asthma are as follows: (1) Children with recurrent wheezing; (2) Wheezing sounds can be heard during wheezing attacks; (3) Positive bronchial dilation test. This project has been approved by the Medical Ethics Committee of Jiangxi Maternal and Child Health Hospital and has agreed to conduct this retrospective study. All patients have obtained informed consent from their guardians and signed informed consent forms, and their clinical information is strictly confidential. In addition, this study strictly adheres to the Helsinki Declaration. The visualization summary of the inclusion process and prediction model construction for pediatric patients is shown in Figure 1.

Liquid Chromatography-Mass Spectrometry (LC-MS)

We extracted 5mL of peripheral blood from the children (10 children with LRTIs and 10 children with secondary asthma caused by LRTIs) included in the study and placed it in a dry test tube without anticoagulants. After standing for 1 hour, the blood was allowed to coagulate and stratify. Then, it was centrifuged for 5 minutes (3000r/min) and the supernatant was taken and centrifuged for another 5 minutes (5000r/min). The supernatant was divided into 2mL centrifuge tubes,

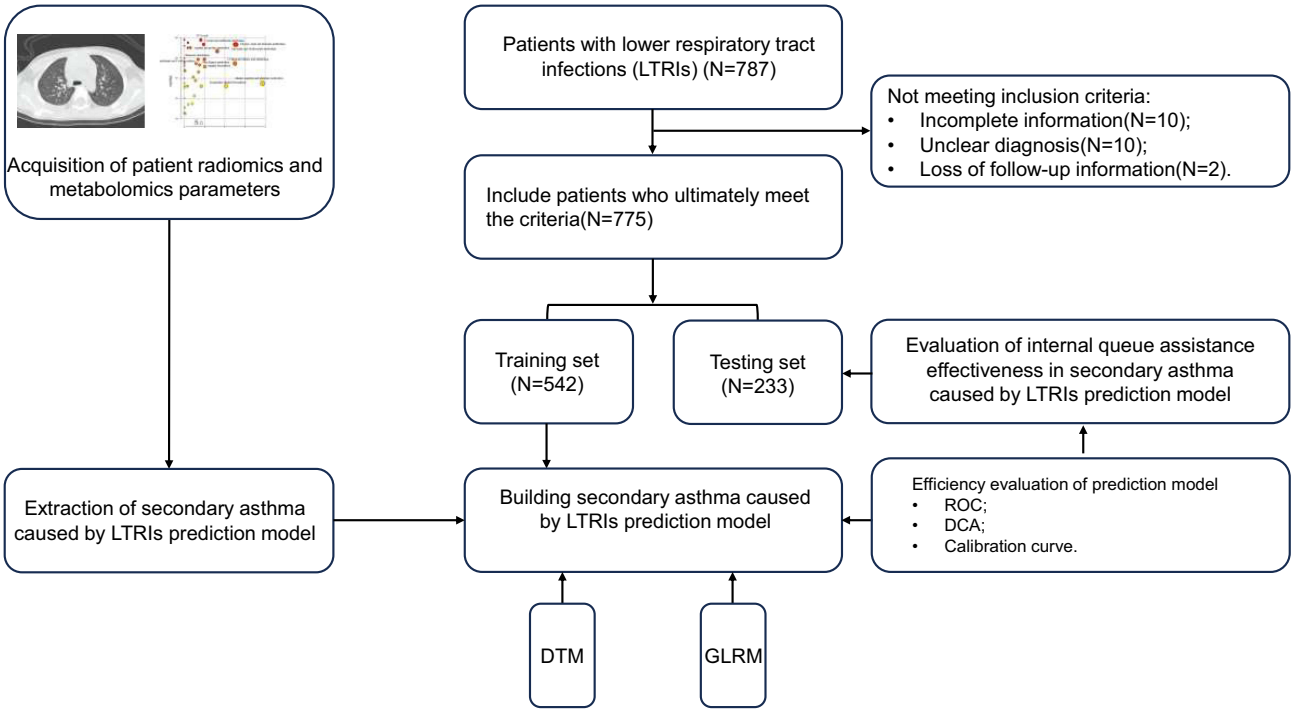

Figure 1 The workflow for screening of pediatric patients and construction of LRTIs secondary asthma prediction model.

with 0.2mL of each tube labeled and stored in a −80 °C freezer. LC-MS detection was performed under the following chromatographic conditions: (1) C18 column, 40 °C, 0.2mL/min; (2) The normal mode is 0.1% formic acid (mobile phase A) and methanol (mobile phase B); Negative mode: 5mM ammonium acetate, pH 9.0 (mobile phase A); (3) Elution procedure: 0.1% formic acid to methanol volume ratio 98:2, 0min → 0.1% formic acid to methanol volume ratio 98:2, 1.5min → 0.1% formic acid to methanol volume ratio 0:100, 3min → 0.1% formic acid to methanol volume ratio 0:100, 10min → 0.1% formic acid to methanol volume ratio 98:2, 10min to 12min, repeated 3 times. Finally, we annotated the identified metabolites using the Small Molecule Pathway Database (SMPDB) and performed principal component analysis on relevant meaningful variables.

CT Image Analysis and Radiomics Feature Extraction

We imported all the chest thin-layer CT images of the children upon admission into 3D Slicer software in DICOM format from the PACS system and used a lung window (window level −500HU, window width 1500HU) to delineate the three-dimensional volume of interest (3D-VOI) of the lungs layer by layer. Next, we used the 3D Slicer amplification software package Pyradiomics to extract radiomic features from suspicious lesion images, including morphological features, gray level co-occurrence matrix, gray level correlation matrix, first-order statistical features, and gray level run length matrix features. Finally, we used Mann Whitney *U*-test to compare the inter-group differences in radiomics, and performed Pearson correlation coefficient analysis on features with statistically significant differences. High redundancy features with $r > 0.8$ were excluded. In order to eliminate the influence of different scales on radiomics features, we standardized the radiomics features using Z-score and used 10-fold cross-validation lasso regression to screen features selected through single-factor analysis. Finally, we repeated the Lasso regression 500 times and selected features that were selected more than 100 times by Lasso regression to establish radiomics signals through multiple logistic regression and shapely additive explanations.[18]

Construction and Validation of Predictive Models

In order to construct a predictive model for LRTIs combined with asthma, we used a nomogram of a generalized linear regression model and a decision tree visualization prediction model. By integrating the clinical baseline data, peripheral

blood metabolomics (LC-MS), and radiomics characteristics of the patient, and comparing and screening the dominant model (as shown in Figure 1 for the construction process of the predictive model). Next, we evaluate the accuracy of the prediction model using the subject's working characteristic curve and related indicators, and evaluate the consistency and calibration performance of the prediction model using a calibration curve. Decision curve analysis evaluates the clinical net benefit values under different threshold probabilities to determine the clinical practicality of multimodal prediction models in metabolomics and radiomics.

Statistical Analysis

All statistical analysis and visualization were completed using R software (version 4.2.3). Independent sample *t*-test, Mann Whitney *U*-test, and chi-square test are used to compare continuous variables and categorical variables. DeLong test, calibration curve, and Hosmer Lemeshow (H-L test) are used to evaluate the predictive performance of the prediction model. A bilateral P value <0.05 is considered statistically significant.

Results

Clinical Demographic and Multi-Omics Parameter Analysis of Children with LRTIs

The distribution characteristics of pathogens in LRTIs showed that 792 strains of pathogens were isolated from 775 children with LRTIs through bacterial culture, including 261 Gram positive bacteria (32.95%) and 531 Gram negative bacteria (67.05%). The drug sensitivity results showed that Staphylococcus aureus had no resistance to vancomycin and had the highest resistance rate to penicillin G, followed by ciprofloxacin and levofloxacin; Streptococcus pneumoniae is not resistant to imipenem and vancomycin but has the highest resistance rate to penicillin G, followed by levofloxacin and benzylpenicillin; Staphylococcus epidermidis is not resistant to imipenem and vancomycin but has the highest resistance rate to benzylpenicillin, followed by penicillin and tetracycline; Escherichia coli is not resistant to imipenem but has the highest resistance to azuron, followed by ampicillin and ceftriaxone; Klebsiella pneumoniae has no resistance to imipenem, no resistance to meropenem, and the highest resistance to ampicillin, followed by ampicillin and cefotaxime; Pseudomonas aeruginosa has no resistance to imipenem and the highest resistance to ampicillin, followed by cefoperazone and cefuroxime. As shown in Table 1 and Supplementary Table 1, there was no statistically significant difference (P > 0.05) between the asthma group and the non-asthma group in terms of baseline data such as gender, age, body mass index (BMI), mode of delivery (MOD), premature delivery, and feeding method. There was a statistically significant difference (P < 0.05) in the comparison of glycerophospholipids, sphingolipids, and radiomics factors. In addition, through enrichment analysis of differential metabolites in plasma metabolomics, we found that differential metabolites were mainly enriched in Sphingolipid Metabolism, Glycerolipid Metabolism, and Phospholipid Biosynthesis (Supplementary Figure 1).

Multi-Omics Feature Selection and Prediction Model Construction

Based on the candidate variables with statistical differences mentioned above, as shown in Figure 2A, Pearson correlation analysis was conducted to retain features with r values greater than 0.8. The dataset was randomly divided into training and testing sets in a ratio of 7–3. Next, we further screened the features using LASSO and obtained 7 non-zero weight coefficient features with the optimal parameter λ=0.017 (Figure 2B and C and Supplementary Table 2), namely glycerophospholipids, sphingolipids, wavelet_LLH_glcm_Cluster Shade (Feature2), wavelet_LHL_glcm_Cluster Prominence (Feature6), wavelet_HLL_glcm_ldmn (Feature7), wavelet_LHL_glcm Correlation (Feature9), and wavelet_HLH_ glrlm Gray Level Variance (Feature11). In addition, we further confirmed through logistic regression analysis that seven potential candidate variables are independent risk factors for LRTIs combined with asthma (Table 2) and constructed a nomogram based on visual generalized linear regression algorithm (Figure 3A), where the candidate variables are assigned quantitative values, and the child can accept the scores assigned by the quantitative values. The comprehensive score is considered as the risk factor for the child to develop asthma. We also combined the binary discriminant analysis principle of decision trees to assign "node" weights to candidate variables, as shown in Figure 3B. Glycerophospholipids, sphingolipids, Feature2, Feature7, and Feature9 occupy certain weights in the risk prediction of

Table 1 Clinical and Multi-Omics Characteristics of Pediatric Patients in the Training and Testing Cohorts

Variables	Training Cohort			P-Value	Testing Cohort			P-Value
	Overall(N=542)	Yes(N=81)	No(N=461)		Overall(N=233)	Yes(N=35)	No(N=198)	
Sex (%)								
Male	276 (50.9)	38 (46.9)	238 (51.6)	0.508	98 (42.1)	12 (34.3)	86 (43.4)	0.409
Female	266 (49.1)	43 (53.1)	223 (48.4)		135 (57.9)	23 (65.7)	112 (56.6)	
Age (%),year								
≤5	264 (48.7)	38 (46.9)	226 (49.0)	0.818	107 (45.9)	17 (48.6)	90 (45.5)	0.875
>5	278 (51.3)	43 (53.1)	235 (51.0)		126 (54.1)	18 (51.4)	108 (54.5)	
COD (median [IQR]),days	10.00 [7.00, 12.00]	9.00 [7.00, 12.00]	10.00 [7.00, 12.00]	0.999	9.00 [7.00, 12.00]	9.00 [7.00, 12.00]	9.00 [6.25, 12.00]	0.682
BMI (median [IQR]),kg/m2	10.26 [9.16, 11.27]	10.04 [9.01, 10.88]	10.31 [9.20, 11.33]	0.117	10.27 [9.17, 11.38]	11.10 [9.65, 11.92]	10.21 [9.16, 11.29]	0.017
PD (%)								
Yes	258 (47.6)	40 (49.4)	218 (47.3)	0.82	107 (45.9)	12 (34.3)	95 (48.0)	0.189
No	284 (52.4)	41 (50.6)	243 (52.7)		126 (54.1)	23 (65.7)	103 (52.0)	
MOD (%)								
Vaginal delivery	267 (49.3)	35 (43.2)	232 (50.3)	0.289	108 (46.4)	9 (25.7)	99 (50.0)	0.013
Cesarean section	275 (50.7)	46 (56.8)	229 (49.7)		125 (53.6)	26 (74.3)	99 (50.0)	
FP (%)								
Breast-feeding	280 (51.7)	45 (55.6)	235 (51.0)	0.522	119 (51.1)	19 (54.3)	100 (50.5)	0.819
Artificial or assisted feeding	262 (48.3)	36 (44.4)	226 (49.0)		114 (48.9)	16 (45.7)	98 (49.5)	
Smoking (%)								
Yes	281 (51.8)	44 (54.3)	237 (51.4)	0.717	108 (46.4)	15 (42.9)	93 (47.0)	0.79
No	261 (48.2)	37 (45.7)	224 (48.6)		125 (53.6)	20 (57.1)	105 (53.0)	
HOA (%)								
Yes	234 (43.2)	62 (76.5)	172 (37.3)	<0.001	100 (42.9)	24 (68.6)	76 (38.4)	0.002
No	308 (56.8)	19 (23.5)	289 (62.7)		133 (57.1)	11 (31.4)	122 (61.6)	
Allergy (%)								
Yes	273 (50.4)	42 (51.9)	231 (50.1)	0.866	119 (51.1)	14 (40.0)	105 (53.0)	0.216
No	269 (49.6)	39 (48.1)	230 (49.9)		114 (48.9)	21 (60.0)	93 (47.0)	
Glycerophospholipids (median [IQR]),ng/mL	0.88 [0.74, 1.06]	2.43 [1.90, 3.00]	0.84 [0.71, 0.98]	<0.001	0.90 [0.70, 1.08]	2.48 [2.05, 3.14]	0.84 [0.68, 1.02]	<0.001
Sphingolipids (median [IQR]),ng/mL	0.60 [0.47, 0.72]	1.38 [1.13, 1.69]	0.56 [0.44, 0.67]	<0.001	0.61 [0.50, 0.73]	1.27 [0.98, 1.56]	0.58 [0.48, 0.68]	<0.001
Feature1 (median [IQR])	28.30 [19.40, 37.40]	31.30 [22.50, 39.60]	27.70 [19.10, 37.00]	0.032	28.00 [18.20, 38.00]	37.70 [21.15, 41.25]	26.85 [17.83, 36.17]	0.014
Feature2 (median [IQR])	161.00 [121.00, 191.00]	89.00 [78.00, 99.00]	170.00 [141.00, 195.00]	<0.001	159.00 [119.00, 188.00]	82.00 [67.50, 94.00]	166.00 [139.25, 193.00]	<0.001
Feature3 (median [IQR])	3.40 [2.33, 4.39]	3.25 [1.95, 4.25]	3.46 [2.41, 4.39]	0.095	3.32 [2.21, 4.48]	3.32 [2.68, 4.39]	3.32 [2.20, 4.47]	0.807
Feature4 (median [IQR])	3.52 [2.51, 4.47]	3.50 [2.30, 4.52]	3.53 [2.57, 4.45]	0.555	3.36 [2.25, 4.41]	3.61 [2.20, 4.64]	3.33 [2.27, 4.34]	0.446
Feature5 (median [IQR])	32.50 [22.33, 43.88]	34.10 [21.70, 44.00]	32.30 [22.60, 43.80]	0.748	34.20 [24.10, 44.70]	30.70 [23.90, 44.85]	35.30 [24.22, 44.65]	0.62
Feature6 (median [IQR])	0.96 [0.89, 1.04]	2.38 [1.83, 2.76]	0.94 [0.87, 1.00]	<0.001	0.98 [0.89, 1.04]	2.42 [1.69, 2.88]	0.95 [0.88, 1.01]	<0.001
Feature7 (median [IQR])	9.60 [8.90, 10.40]	21.90 [17.80, 26.80]	9.40 [8.80, 10.00]	<0.001	9.50 [8.90, 10.40]	23.20 [18.45, 27.50]	9.40 [8.90, 10.00]	<0.001

(*Continued*)

https://doi.org/10.2147/IJGM.S498965

Table 1 (Continued).

Variables	Training Cohort			P-Value	Testing Cohort			P-Value
	Overall(N=542)	Yes(N=81)	No(N=461)		Overall(N=233)	Yes(N=35)	No(N=198)	
Feature8 (median [IQR])	1.40 [0.98, 1.80]	1.30 [0.94, 1.85]	1.40 [1.01, 1.78]	0.86	1.32 [0.94, 1.71]	1.38 [1.13, 1.86]	1.29 [0.89, 1.68]	0.052
Feature9 (median [IQR])	1.19 [1.01, 1.39]	0.47 [0.32, 0.70]	1.25 [1.10, 1.41]	<0.001	1.16 [0.99, 1.38]	0.50 [0.28, 0.66]	1.25 [1.06, 1.41]	<0.001
Feature10 (median [IQR])	66.10 [43.65, 89.80]	70.90 [44.10, 91.30]	66.10 [43.30, 89.80]	0.732	68.70 [43.80, 89.20]	67.20 [33.70, 78.50]	69.15 [43.97, 89.52]	0.354
Feature11 (median [IQR])	1.56 [1.32, 1.89]	0.91 [0.81, 1.02]	1.66 [1.43, 1.93]	<0.001	1.57 [1.33, 1.85]	0.98 [0.79, 1.08]	1.63 [1.45, 1.89]	<0.001
Feature12 (median [IQR])	0.86 [0.47, 1.24]	0.89 [0.44, 1.21]	0.86 [0.49, 1.25]	0.811	0.85 [0.45, 1.23]	0.79 [0.50, 1.18]	0.85 [0.44, 1.23]	0.903

Abbreviations: COD, course of disease; BMI, body mass index; PD, premature delivery; MOD, mode of delivery; FP, feeding pattern; Smoking, smoking history of parents; HOA, history of asthma; Allergy, history of allergy; FA, fatty Acyls; SSD, steroids and steroid derivatives; OC, organooxygen compounds; PL, prenol lipids; UH, unsaturated hydrocarbons; OPAD, organic phosphoric acids and derivatives; UM, unrecognized metabolites; Feature1, wavelet_LLL_glcm_Cluster Prominence; Feature2, wavelet_LLH_glcm_Cluster Shade; Feature3, wavelet_LHH_ glszm_Gray Leve Variance; Feature4, wavelet_LLH_firstorder_Variance; Feature5, wavelet_LHH_glszm_LowGrayLevelZoneEmphasis; Feature6, wavelet_LHL_glcm_Cluster Prominence; Feature7, wavelet_HLL_glcm_Idmn; Feature8, wavelet_HLH_glszm_Smal Area High Gray Level Emphasis; Feature9, wavelet_LHL_glcm Correlation; Feature10, wavelet_HLH_glrlm_Low Gray Level Run Emphasis; Feature11, wavelet_HLH_ glrlm Gray Level Variance; Feature12, wavelet_HLH_glszm_Large Area Low Gray Level Emphasis.

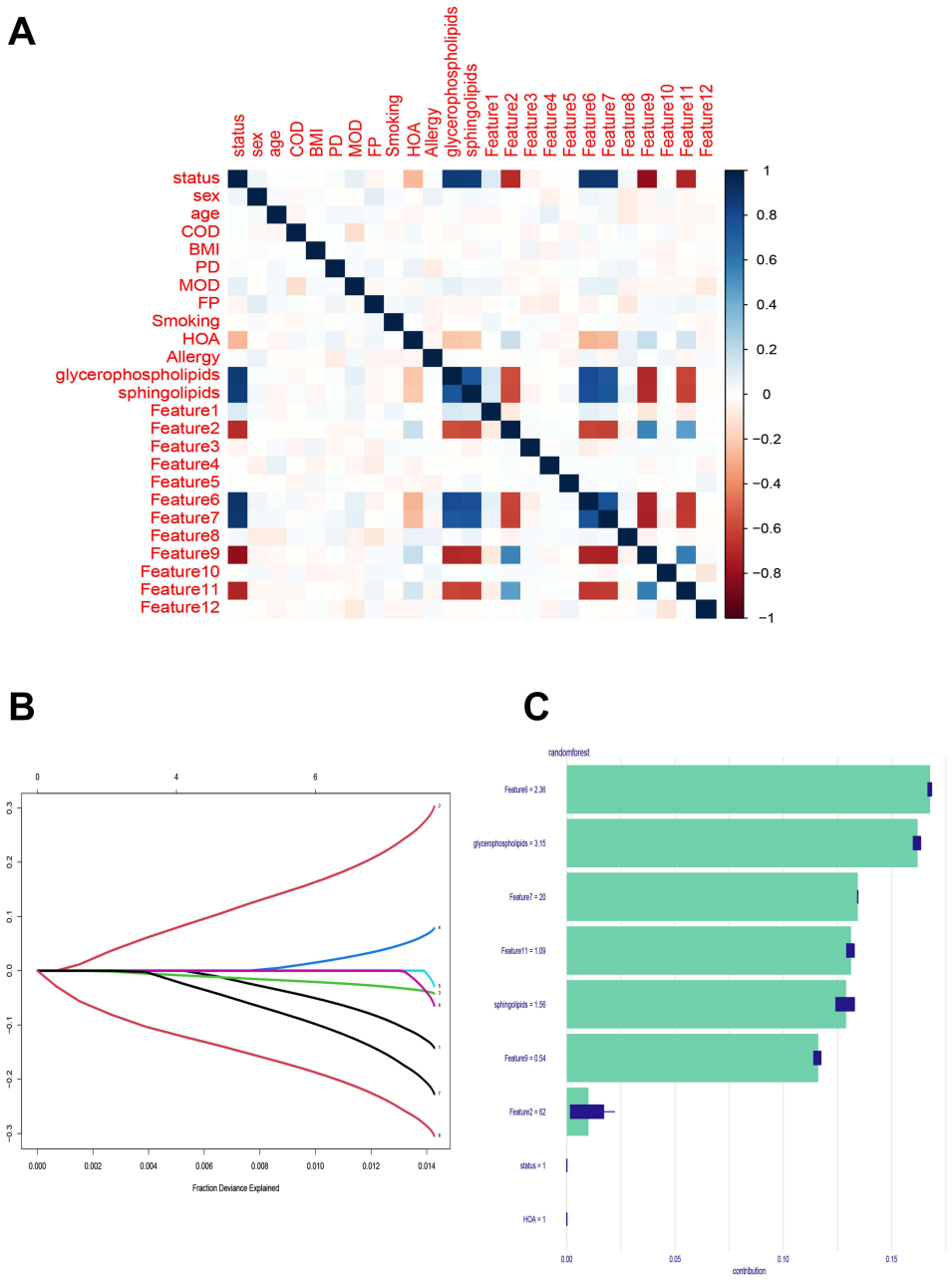

Figure 2 Selected features of multi-omics model. (**A**) Heatmap describing the correlation coefficient matrices of seven selected radiomics features; (**B**) Screening of optimal combination candidate predictor variables based on Lasso algorithm; (**C**) Shapley value method displays the weight allocation of candidate variables.

LRTIs combined with asthma, indicating that the LRTIs combined with asthma prediction model based on multi-omics has visual advantages and clinical operability.

Evaluation of the Efficacy of LRTIs Combined with Asthma Prediction Models

As shown in Table 3, in the prediction model constructed based on the generalized linear regression algorithm, the AUC value of the ROC curve in the training set is 0.817 (0.760–0.874), and the AUC value in the test set is 0.799 (0.742–0.856); The AUC values of the prediction model based on decision trees in the training and testing sets are 0.926 (0.869–0.983) and 0.917 (0.860–0.974), respectively. As for the robustness evaluation of the two prediction models, we compared them using the C-Index in the calibration curve mode. The C-Index values of the generalized linear regression prediction model (ie nomogram) in the training and testing sets were 0.799 and 0.872, respectively. The

Table 2 Univariate and Multivariable Logistic Regression Analyses for Selecting Candidate Predictive Features

Variables	Univariate Analysis		P-Value	Multivariate Analysis		P-Value
	OR	95% CI		OR	95% CI	
HOA	1.23	0.87–2.26	<0.05	1.29	0.77–2.41	0.12
Glycerophospholipids	1.09	0.72–2.14	<0.05	1.11	0.83–2.08	<0.01
Sphingolipids	2.03	1.02–4.79	<0.01	1.98	0.99–3.76	<0.01
Feature2	1.17	0.63–2.26	<0.01	1.14	0.71–2.81	<0.01
Feature6	2.11	0.89–3.42	<0.01	2.08	0.71–3.89	<0.01
Feature7	1.58	0.25–3.44	<0.01	1.61	0.35–3.71	<0.01
Feature9	1.67	0.82–2.73	<0.01	1.63	0.62–2.82	<0.01
Feature11	1.43	0.56–2.89	<0.01	1.52	0.61–2.87	<0.01

Abbreviations: OR, odds ratio; 95% CI, 95% confidence interval; HOA, history of asthma; Feature2, wavelet_LLH_glcm_Cluster Shade; Feature6, wavelet_LHL_glcm_Cluster Prominence; Feature7, wavelet_HLL_glcm_Idmn; Feature9, wavelet_LHL_glcm Correlation; Feature11, wavelet_HLH_ glrlm Gray Level Variance.

C-index values of the decision tree prediction model in the training and testing sets are 0.767 and 0.883, respectively. Consistent with the calibration curve results, we found that decision curve analysis showed that decision trees were more robust and accurate than nomograms in predicting the performance and net benefits of LRTIs combined with asthma (Figure 4). The above results indicate that although LRTIs combined with asthma prediction models constructed based on different machine learning algorithms can distinguish the risk of asthma occurrence, the prediction model constructed by combining decision trees with multi-omics has better performance and is therefore more suitable for clinical decision-making assistance.

Discussion

The etiology and pathogenesis of pediatric respiratory infections combined with asthma have always been very complex, making it an urgent research hotspot in clinical practice.[19] Due to the unclear etiology and triggering factors, as well as the fact that children often exhibit sudden and recurrent characteristics, it poses a great challenge to the diagnosis and treatment of clinical physicians. Previous studies have shown that pediatric asthma is essentially limited to chronic non-specific changes in the airway and airway hyperresponsiveness.[20,21] However, there is insufficient understanding of lower respiratory tract infections as an independent risk factor, resulting in a high mortality rate in children diagnosed with LRTIs combined with asthma in the early stages. In view of this, this study analyzed for the first time the risk factors of LRTIs combined with asthma in children, innovatively used multi-omics analysis to screen the molecular mechanisms and potential biomarkers associated with LRTIs combined with asthma, and constructed a multimodal machine learning prediction model to guide clinical decision-making, especially for early identification of children with high-risk LRTIs combined with asthma, which has crucial guiding significance.

Previous studies have shown that metabolic disorders may be involved in the pathogenesis of pulmonary fibrosis, but the association between asthma pathogenesis and metabolomics has not been fully revealed.[22–24] This study shows that differential metabolites are mostly glycerophospholipids, which are important components of cell membranes and can participate in mediating airway inflammation through phospholipase A2 to produce lysophosphatides, thereby participating in signal transduction and immune responses. Recent studies have shown that glycerophospholipids are involved in lung infections, asthma, and chronic obstructive pulmonary disease, while lipids play a crucial role in the lungs as a structural component, energy storage, and surfactant molecular signaling mediator.[25,26] As a broad class of compounds, sphingolipids mainly originate from 18 carbon amino alcohols, which are synthesized in the endoplasmic reticulum and participate in various pathological and physiological processes within cells, such as cell structure, storage, and signal transduction.[27,28] At present, there is evidence to suggest that these molecules are key mediators of immune response, especially sphingolipids, which have become potential regulators of pulmonary fibrosis, as they can regulate cell migration, promote apoptosis, and facilitate epithelial and endothelial cell cycle arrest.[29–31] Our study also suggests

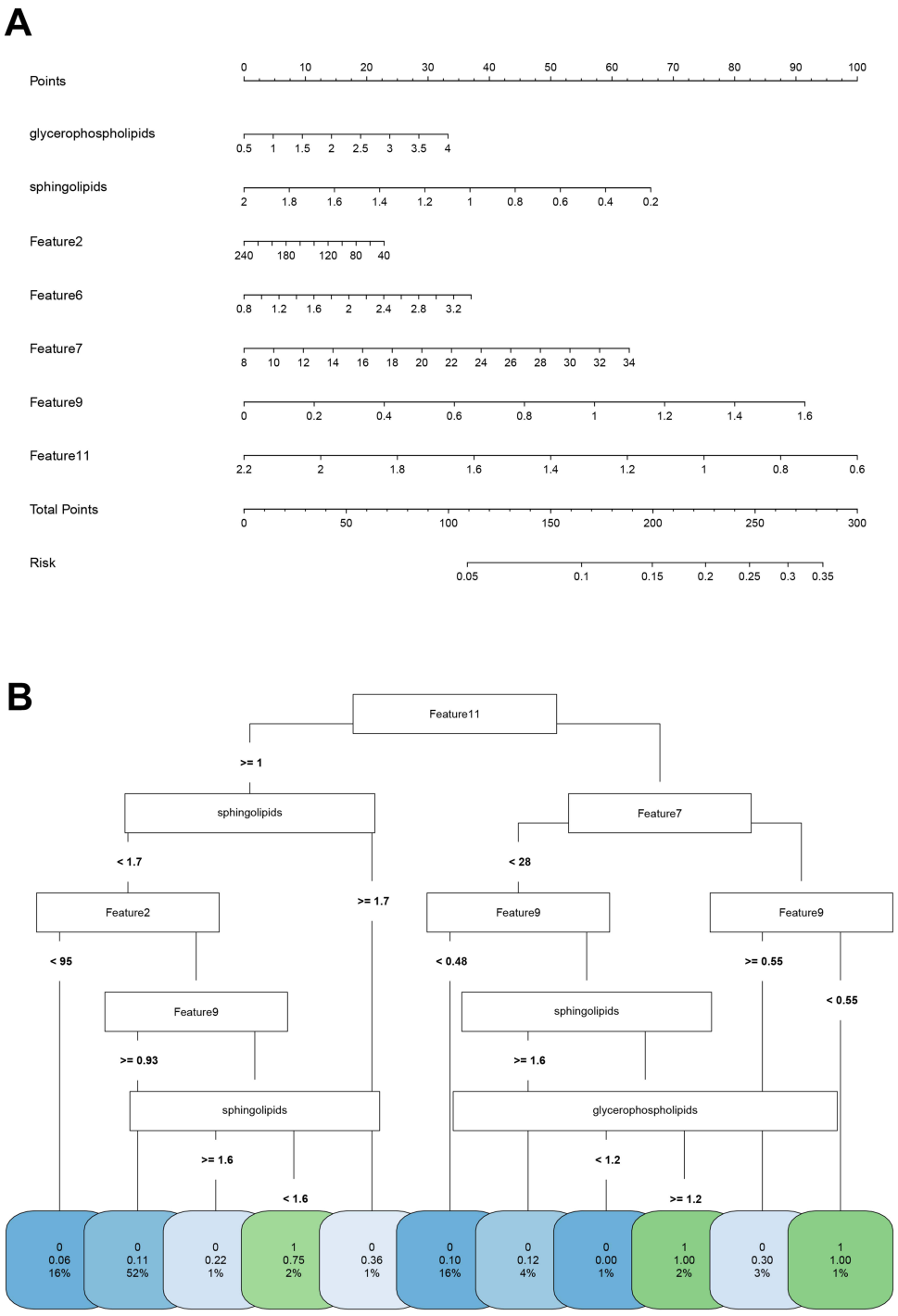

Figure 3 Constructing a visual prediction model for secondary asthma caused by LTRIs based on multi omics.(**A**) Nomogram;(**B**) Decision tree model.
Notes: As for the nomogram, each predictor variable is assigned a value, and the risk of developing asthma in children can be considered as the risk coefficient for secondary asthma in LTRIs based on the total value assigned to each variable, ie the risk value corresponding to the "Risk" total score. Each step node of the decision tree is assigned candidate variables, which are selected one by one based on binary discriminant analysis of the candidate variables. The final outcome is the risk outcome of secondary asthma in LTRIs.

that ceramides as predictive markers for LRTIs combined with asthma may affect the progression of LRTIs disease and participate in the pathophysiological processes of asthma. Therefore, the next step is to further explore the early biomarkers and potential molecular mechanisms of LRTIs combined with asthma using lipidomics.

Table 3 Predictive Performance of Predictive Models in the Training and Testing Cohort

Prediction Model	Training Set				International Set			
	AUC	95% CI	PPV	NPV	AUC	95% CI	PPV	NPV
DTM	0.926	0.869–0.983	0.967	0.878	0.917	0.860–0.974	0.991	0.803
GLRM	0.817	0.760–0.874	0.823	0.766	0.799	0.742–0.856	0.816	0.786

Abbreviations: AUC, Area under the curve;95% CI, 95% confidence interval;PPV, Positive predictive value;NPV, negative predictive value; DTM, Decision tree model; GLRM, Generalized linear regression.

Radiomics can extract image information features for high-throughput evaluation of the intrinsic heterogeneity and invasiveness of tumors and has been widely used in the differentiation of benign and malignant lung nodules, gene mutation prediction, and prognosis assessment.[32,33] Nowadays, medical imaging has always been an effective tool for evaluating asthma, which can visualize the lungs of patients, evaluate local airway characteristics, and specifically display the location and cause of airway obstruction in asthma patients. However, there have been few reports in previous literature on radiomics studies of LRTIs, especially when LRTIs are associated with asthma. The traditional concept holds that asthma is mainly characterized by airway lesions and does not involve lung parenchyma.[34] However, in this study, texture features of lesions were extracted from the CT lung window of pediatric patients, and a total of 156 usable feature parameters were extracted. Finally, 12 texture features with predictive significance were selected, covering first-order to higher-order texture features, which can better reflect the spatial properties of lesions. The results showed that the AUC of the decision tree prediction model based on texture features combined with metabolomics for diagnosing LRTIs combined with asthma in the training set and validation set were 0.926 and 0.917, respectively. This provides us with innovative insights that imaging indicators related to airway remodeling, bronchiectasis, and mucus embolism in children with asthma can serve as novel predictive indicators. Combined with clinical indicators, these indicators can be used to predict lung function and treatment responsiveness in children with LRTIs and asthma, providing new ideas for individualized treatment of LRTIs combined with asthma.

In this study, we also attempted to use machine learning algorithms to evaluate the predictive performance of different prediction models, which brought us new insights. The LRTIs merged asthma prediction model constructed using decision tree algorithm has better predictive performance, while the traditional generalized linear regression prediction model has significantly lower predictive performance than decision trees. Consistent with previous research results, decision trees, as a branch of random forest discriminant analysis, strictly follow the binary classification discriminant iterative analysis process in their decision-making process.[35,36] In this study, we utilized radiomics and metabolomics candidate variables to achieve initial screening of high-risk asthma children, and their predictive value is self-evident. However, even if the same variable is applied to prediction models of different algorithms, its weights and effect values are not the same. Therefore, combining multiple candidate indicators and relying on advanced algorithms to construct LRTIs combined asthma prediction models has extremely important guiding value for clinical sorting of high-risk children. However, the extended training of the two common prediction models mentioned above still needs to be validated and optimized in the future.

This study also inevitably has the following limitations. First, as a single-center retrospective study, the patient cohort is bound to have geographical limitations and selection bias. Although we used multiple corrections and repeated sampling, as well as internal validation in our analysis, we still need multi-center, large sample prospective study cohorts for model expansion training in the future; Secondly, the combination of metabolomics and radiomics candidate parameters was used to screen for predictable LRTIs combined with asthma predictive factors. Although seven candidate predictive parameters have been screened, their clinical combined or single use predictive performance, as well as whether metabolic factors are affected by geography and detection technology, still need to be further validated. We plan to invest more funds to focus on valuable metabolic factors and conduct in-depth research on candidate biomarkers and pathogenesis; Thirdly, conventional machine learning algorithms such as generalized linear regression and decision trees are included in this study. In the future, more advanced algorithms (such as artificial neural networks, support vector

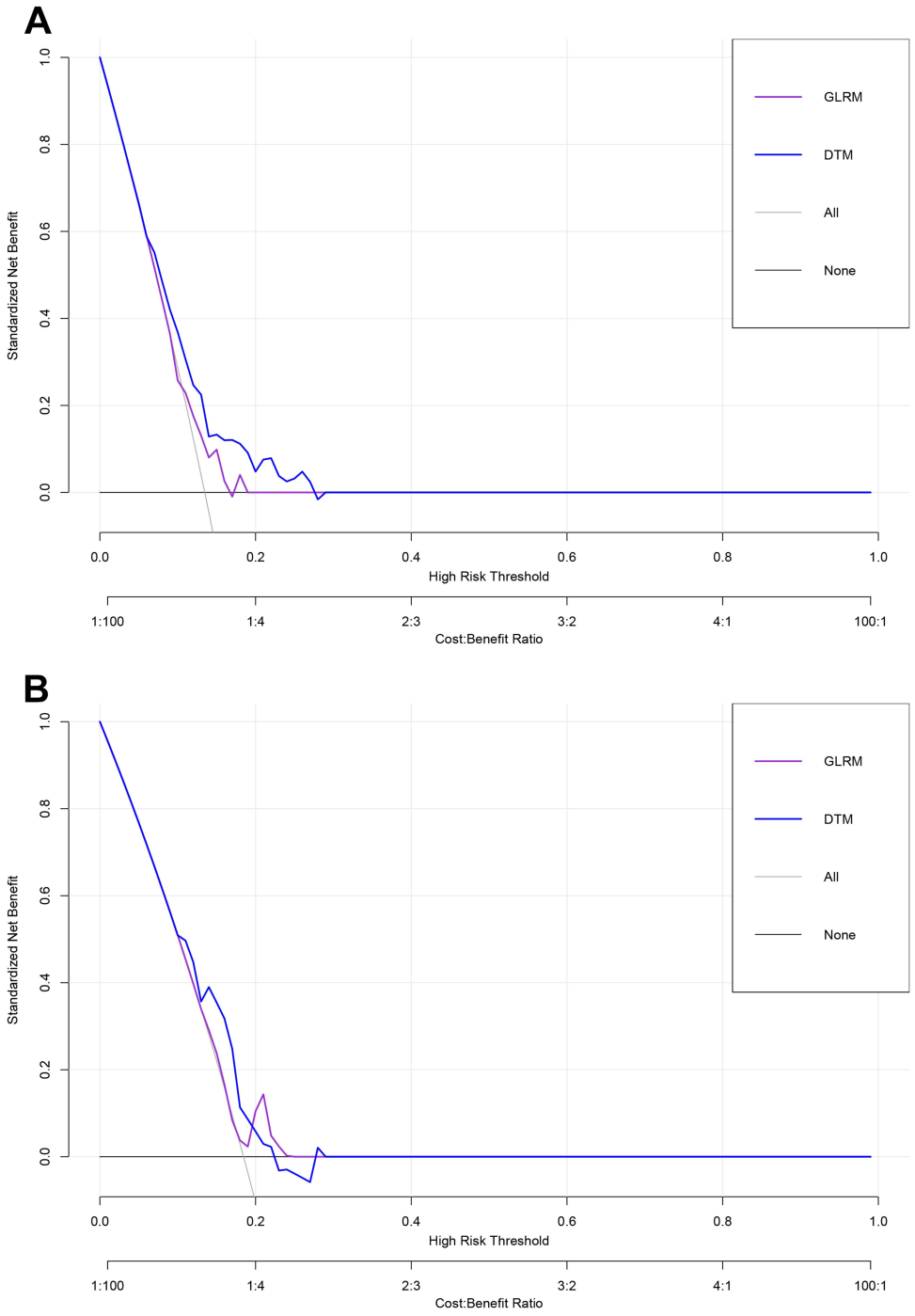

Figure 4 Decision curve analysis (DCA) for two predictive models in the training and testing cohort. (**A**) Training cohort; (**B**) Testing cohort. The curves indicated that the decision tree model had a higher clinical benefit when compared to the GLR model, with a threshold probability range of 0.35 to 0.9.

machines, naive Bayes, etc.) need to be applied to the training of prediction models in order to find better prediction models and assist in the clinical decision-making of LRTIs combined with asthma.

Conclusion

In summary, a combination model that combines clinical radiological features with metabolomics can be an effective strategy for diagnosing LRTIs combined with asthma, especially the new fusion multimodal omics prediction model

based on decision trees, which will help provide decision support for early identification and treatment planning of high-risk asthma in LRTIs patients.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Claassen-Weitz S, Lim KYL, Mullally C, Zar HJ, Nicol MP. The association between bacteria colonizing the upper respiratory tract and lower respiratory tract infection in young children: a systematic review and meta-analysis. *Clin Microbiol Infect*. 2021;27(9):1262–1270. doi:10.1016/j.cmi.2021.05.034
2. Klig JE, Chen L. Lower respiratory infections in children. *Curr Opin Pediatr*. 2003;15(1):121–126. doi:10.1097/00008480-200302000-00020
3. Troeger C, Forouzanfar M, Rao PC. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Infect Dis*. 2017;17(11):1133–1161. doi:10.1016/S1473-3099(17)30396-1
4. Collaro AJ, McElrea MS, Marchant JM, et al. The effect of early childhood respiratory infections and pneumonia on lifelong lung function: a systematic review. *Lancet Child Adolesc Health*. 2023;7(6):429–440. doi:10.1016/S2352-4642(23)00030-5
5. Liu Y, Liu J, Chen F, et al. Impact of meteorological factors on lower respiratory tract infections in children. *J Int Med Res*. 2016;44(1):30–41. doi:10.1177/0300060515586007
6. Lowe MC. Childhood Respiratory Conditions: lower Respiratory Tract Infection. *FP Essentials*. 2022;513:20–24.
7. Madhi F, Panetta L, De Pontual L, et al. Antimicrobial treatment of lower respiratory tract infections in children. *Infect Dis*. 2023;53(8s):104782. doi:10.1016/j.idnow.2023.104782
8. Cohen R, Angoulvant F, Biscardi S, Madhi F, Dubos F, Gillet Y. Antibiotic treatment of lower respiratory tract infections. *Archiv Pediatrie*. 2017;24(12s):S17–s21. doi:10.1016/S0929-693X(17)30513-4
9. Onakpoya IJ, Hayward G, Heneghan CJ. Antibiotics for preventing lower respiratory tract infections in high-risk children aged 12 years and under. *Cochrane Database Syst Rev*. 2015;2015(9):Cd011530. doi:10.1002/14651858.CD011530.pub2
10. Brittan MS, Moss A, Watson JD, et al. Association between early childhood lower respiratory tract infections and subsequent asthma. *J Asthma*. 2022;59(11):2143–2153. doi:10.1080/02770903.2021.1999469
11. Geerdink RJ, Pillay J, Meyaard L, Bont L. Neutrophils in respiratory syncytial virus infection: a target for asthma prevention. *J Allergy Clin Immunol*. 2015;136(4):838–847. doi:10.1016/j.jaci.2015.06.034
12. Huang AA, Huang SY, Gholamzadeh M. Use of feature importance statistics to accurately predict asthma attacks using machine learning: a cross-sectional cohort study of the US population. *PLoS One*. 2023;18(11):e0288903. doi:10.1371/journal.pone.0288903
13. Molfino NA, Turcatel G, Riskin D. Machine Learning Approaches to Predict Asthma Exacerbations: a Narrative Review. *Advances in Therapy*. 2024;41(2):534–552. doi:10.1007/s12325-023-02743-3
14. Babu M, Snyder M. Multi-Omics Profiling for Health. *Mol Cellular Proteom*. 2023;22(6):100561. doi:10.1016/j.mcpro.2023.100561
15. Wang L, Qu J, Harari O, Boddey JA, Wang Z, Linna-Kuosmanen S. The impact of multi-omics in medicine. *Cell Rep Med*. 2024;5(9):101742. doi:10.1016/j.xcrm.2024.101742
16. Dondi F, Gatta R, Treglia G, et al. Application of radiomics and machine learning to thyroid diseases in nuclear medicine: a systematic review. *Rev Endocr Metab Disord*. 2024;25(1):175–186. doi:10.1007/s11154-023-09822-4
17. Fujiogi M, Zhu Z, Raita Y, et al. Nasopharyngeal lipidomic endotypes of infants with bronchiolitis and risk of childhood asthma: a multicentre prospective study. *Thorax*. 2022;77(11):1059–1069. doi:10.1136/thorax-2022-219016
18. Huang AA, Huang SY, Bellantuono L. Increasing transparency in machine learning through bootstrap simulation and shapely additive explanations. *PLoS One*. 2023;18(2):e0281922. doi:10.1371/journal.pone.0281922
19. Liu WJ, Lo WS, Hsieh HJ, Huang TY, Chiang CH. Increased incidence of asthma and pulmonary dysfunction after severe lower respiratory tract infection in infancy. *Zhonghua Minguo xiao er ke yi xue hui za zhi [Journal] Zhonghua Minguo xiao er ke yi xue hui. Zhonghua Minguo xiao er ke yi xue hui za zhi [Journal]. Zhonghua Minguo xiao er ke yi xue hui*. 1991;32(6):348–357.
20. de Benedictis FM, Attanasi M, de Benedictis FM. Asthma in childhood. *Eur Respirat Rev*. 2016;25(139):41–47. doi:10.1183/16000617.0082-2015
21. Liapikou A, Torres A. Pharmacotherapy for lower respiratory tract infections. *Expert Opinion Pharmacother*. 2014;15(16):2307–2318. doi:10.1517/14656566.2014.959927
22. Chung MJ, Lee KS, Franquet T, Müller NL, Han J, Kwon OJ. Metabolic lung disease: imaging and histopathologic findings. *Eur J Radiol*. 2005;54(2):233–245. doi:10.1016/j.ejrad.2004.07.003
23. Macklin M, Thompson C, Kawano-Dourado L, et al. Linking Adiposity to Interstitial Lung Disease: the Role of the Dysfunctional Adipocyte and Inflammation. *Cells*. 2023;12(18):2206. doi:10.3390/cells12182206
24. Georgakopoulou VE, Lempesis IG, Sklapani P, Trakas N, Spandidos DA. Precision medicine for respiratory diseases: a current viewpoint. *Med Int*. 2024;4(4):31. doi:10.3892/mi.2024.155
25. Choi IW, Sun K, Kim YS, et al. TNF-alpha induces the late-phase airway hyperresponsiveness and airway inflammation through cytosolic phospholipase A(2) activation. *J Allergy Clin Immunol*. 2005;116(3):537–543. doi:10.1016/j.jaci.2005.05.034
26. Ko HM, Kang NI, Kim YS, et al. Glutamine preferentially inhibits T-helper type 2 cell-mediated airway inflammation and late airway hyperresponsiveness through the inhibition of cytosolic phospholipase A(2) activity in a murine asthma model. *Clin Experiment Allergy*. 2008;38(2):357–364. doi:10.1111/j.1365-2222.2007.02900.x
27. Hannun YA, Obeid LM. Sphingolipids and their metabolism in physiology and disease. *Nat Rev Mol Cell Biol*. 2018;19(3):175–191. doi:10.1038/nrm.2017.107
28. Green CD, Maceyka M, Cowart LA, Spiegel S. Sphingolipids in metabolic disease: the good, the bad, and the unknown. *Cell Metab*. 2021;33(7):1293–1306. doi:10.1016/j.cmet.2021.06.006

29. Huang LS, Natarajan V. Sphingolipids in pulmonary fibrosis. *Adv Bio Regul*. 2015;57:55–63. doi:10.1016/j.jbior.2014.09.008
30. Jayant G, Kuperberg S, Somnay K, Wadgaonkar R. The Role of Sphingolipids in Regulating Vascular Permeability in Idiopathic Pulmonary Fibrosis. *Biomedicines*. 2023;11(6):1728. doi:10.3390/biomedicines11061728
31. Aureli M, Schiumarini D, Loberto N, et al. Unravelling the role of sphingolipids in cystic fibrosis lung disease. *Chem Phys Lipids*. 2016;200:94–103. doi:10.1016/j.chemphyslip.2016.08.002
32. Avanzo M, Stancanello J, Pirrone G, Sartor G. Radiomics and deep learning in lung cancer. *Strahlentherapie und Onkologie*. 2020;196 (10):879–887. doi:10.1007/s00066-020-01625-9
33. Fornacon-Wood I, Faivre-Finn C, Jpb O, Price GJ. Radiomics as a personalized medicine tool in lung cancer: separating the hope from the hype. *Lung Cancer*. 2020;146:197–208. doi:10.1016/j.lungcan.2020.05.028
34. Barrios RJ, Kheradmand F, Batts L, Corry DB. Asthma: pathology and pathophysiology. *Arch Pathol Lab Med*. 2006;130(4):447–451. doi:10.5858/2006-130-447-APAP
35. Hammann F, Drewe J. Decision tree models for data mining in hit discovery. *Expert Opin Drug Discov*. 2012;7(4):341–352. doi:10.1517/17460441.2012.668182
36. Sherafatian M, Arjmand F. Decision tree-based classifiers for lung cancer diagnosis and subtyping using TCGA miRNA expression data. *Oncol Lett*. 2019;18(2):2125–2131. doi:10.3892/ol.2019.10462