Chen et al. World Allergy Organization Journal (2023) 16:100796
http://doi.org/10.1016/j.waojou.2023.100796

WORLD ALLERGY ORGANIZATION JOURNAL

Open Access

A novel inflammatory endotype diagnostic model based on cytokines in chronic rhinosinusitis with nasal polyps

Mengyu Chen, MD[a,b,c,1], Zhaofeng Xu, MD[a,b,c,1], Yiwei Fu, MD[d,g,1], Nan Zhang, MD, PhD[a,b,c,e], Tong Lu, MD, PhD[a,b,c], Zhengqi Li, MD, PhD[a,b,c], Jian Li, MD, PhD[a,b,c], Claus Bachert, MD, PhD[a,b,c,e***], Weiping Wen, MD, PhD[a,b,c,f**] and Yihui Wen, MD, PhD[a,b,c*]

ABSTRACT

Background: Type 2 CRSwNP is characterized by severe symptoms, multiple comorbidities, longer recovery course and high recurrence rate. A simple and cost-effective diagnostic model for CRSwNP endotype integrating clinical characteristics and histopathological features is urgently needed.

Objective: To establish a clinical diagnostic model of inflammatory endotype in CRSwNP based on the clinical characteristics, pathological characteristics, and cytokines profile in the polyp tissue of patients.

Methods: A total of 244 participants with CRSwNP were enrolled at 2 different centers in China and Belgium from 2018 to 2020. IL-5 level of nasal polyp tissue was used as gold standard. Clinical characteristics were used to establish diagnostic models. The area under the receiver operating curve (AUC) was used to evaluate the diagnostic performance. The study was approved by the ethics board of the First Affiliated Hospital of Sun Yat-sen University ([2020] 302), and written informed consent was obtained from all subjects before inclusion.

Results: In total, 134 patients from China (training set) and 110 patients from Belgium (validation set) were included. The logistic regression (LR) model in predicting inflammatory endotype of CRSwNP showed the AUC of 83%, which was better than the diagnostic performance of machine learning models (AUC of 61.14%–82.42%), and single clinical variables. We developed a simplified scoring system based on LR model which shows similar diagnostic performance to the LR model ($P = 0.6633$).

Conclusion: The LR model in this diagnostic study provided greater accuracy in prediction of inflammatory endotype of CRSwNP than those obtained from the machine learning model and single clinical variable. This indicates great potential for the use of diagnostic model to facilitate inflammatory endotype evaluation when tissue cytokines are unable to be measured.

[a]Department of Otolaryngology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong, PR China
*Corresponding author. Department of Otolaryngology, The First Affiliated Hospital, Sun Yat-sen University, No 58, Zhong Shan Er Lu, Guangzhou, 510000, Guangdong, PR China. E-mail: wenyihui@mail.sysu.edu.cn
**Corresponding author. Department of Otolaryngology, The First Affiliated Hospital, Sun Yat-sen University, No 58, Zhong Shan Er Lu, Guangzhou 510000, Guangdong, PR China. Email: wenwp@mail.sysu.edu.cn
***Corresponding author. The First Affiliated Hospital, Sun Yat-sen University, International Airway Research Center, No 58, Zhong Shan Er Lu, Guangzhou, 510000, Guangzhou, Guangdong, PR China. Email: bachert@mail.sysu.edu.cn

[1] These authors have contributed equally to this work.
Full list of author information is available at the end of the article

http://doi.org/10.1016/j.waojou.2023.100796
Received 9 March 2023; Received in revised from 12 June 2023; Accepted 19 June 2023
Online publication date xxx

Keywords: Biomarkers, Chronic rhinosinusitis, Diagnostic model, Inflammatory endotypes, Type-2 inflammation

INTRODUCTION

Chronic rhinosinusitis (CRS) is a common mucosal inflammatory disease of paranasal sinuses with a course of more than 12 weeks. The prevalence of CRS is 12% in Western countries and 8% in China.[1–3] Broadly, CRS is classified into 2 major phenotypes: CRS without nasal polyps (CRSsNP) and CRS with nasal polyps (CRSwNP). CRSwNP is associated with higher levels of morbidity and can influence lower airway disease status in adults, which is considered to be different from CRSsNP in the inflammatory profile.[4]

In clinical studies, the histopathological characteristics and infiltrating inflammatory cell types of CRSwNP were distinct among different populations and regions.[5] However, the eosinophilic endotype of CRSwNP tends to increase in Asia over the past few decades.[6–8] Marked progress has been made in understanding the mucosal immunology of CRSwNP, and classifying CRSwNP endotypes according to the dominant T helper cell response and corresponding cytokine levels of paranasal mucosa is essential for accurate diagnoses and individual treatments. Type 2 CRSwNP is characterized by T helper type 2 cell (Th2) response with the involvement of IL-4, IL-5, and IL-13, resulting in eosinophils infiltration in polyps,[9] while non-type 2 CRSwNP is dominated by other types of T helper cell, such as Th1 and Th17 cells. Therefore, CRSwNP should be regarded as a heterogeneous disease with multiple endotypes. Type 2 CRSwNP is associated with high risks of recurrence and decreased quality of life, making this disease clinically important to identify, evaluate, and treat.

Although identifying inflammatory endotypes by measuring cytokines in nasal polyps (NP) can assist in personalized treatment, this method is not practical for daily clinical work. The disadvantages of high expenses and timelag hinder the promotion in clinical settings. In clinical practice, the degree of eosinophil infiltration in polyp tissue is often used as an alternative to classify CRSwNP endotypes.[10] However, there are no uniform criteria regarding histopathological classification. Using mucosal eosinophilic status as a single indicator to diagnose type 2 CRSwNP was inadequate. Other factors, such as comorbidities, should also be taken into account. A simple and cost-effective diagnostic model integrating clinical characteristics, histopathological features, and cytokine level for CRSwNP endotype is urgently needed.

In this study, we first defined IL-5 level of polyps as the gold standard for CRSwNP endotype diagnosis based on biomarker cluster analysis. Several models were established and evaluated using the clinical and histopathological characteristics of CRSwNP patients. Then we transformed the optimal model into a simple and accurate clinical scoring system to assess the inflammatory endotype of patients with CRSwNP for clinical practice.

METHODS

Participants selection and endotyping

This was a retrospective study involving patients with CRSwNP hospitalized in both China and Belgium from January 2018 to December 2020. Ethics Committee for Clinical Research and animal trials approved the study ([2020]302), and written informed consent was obtained from all subjects before inclusion. The CRSwNP was diagnosed according to European Position Paper on Rhinosinusitis and Nasal Polyps 2020 (EPOS2020).[11] NP tissues from 70 patients at China and 110 patients at Belgium with CRSwNP were collected during surgery. None of the subjects had taken oral corticosteroids, immunomodulatory drugs within 4 weeks, nasal corticosteroids, or oral antibiotics within 2 weeks before surgery. Patients with fungal sinusitis, antrochoanal polyp, gastroesophageal reflux disease, immunodeficiency disease, primary ciliary immobility syndrome, cystic fibrosis, and parasitic

	level	Training set	Validation set	P
n		134	110	
Age (mean (SD))		45.90 (13.52)	48.85 (12.71)	0.083
Gender (%)	Female	57 (42.5)	32 (29.1)	0.042*
	Male	77 (57.5)	78 (70.9)	
Smoke (%)	No	110 (82.1)	41 (60.3)	0.001*
	Yes	24 (17.9)	27 (39.7)	
Alcohol (%)	No	119 (88.8)	90 (88.2)	1
	Yes	15 (11.2)	12 (11.8)	
Anosmia (%)	No	17 (12.7)	12 (10.9)	0.82
	Yes	117 (87.3)	98 (89.1)	
Asthma (%)	No	45 (33.6)	61 (55.5)	0.001*
	Yes	89 (66.4)	49 (44.5)	
NERD (%)	No	130 (97.01)	91 (82.73)	<0.001*
	Yes	4 (2.99)	19 (17.27)	
Recurrence (%)	No	96 (71.6)	29 (27.4)	<0.001*
	Yes	38 (28.4)	77 (72.6)	
EOS count (median [IQR])		0.40 (0.21, 0.57)	0.36 (0.21, 0.57)	0.711
ECRSwNP (%)	No	35 (26.1)	29 (26.4)	1
	Yes	99 (73.9)	81 (73.6)	
Endotype (%)	Non-Type 2	24 (17.9)	21 (19.1)	0.944
	Type 2	110 (82.1)	89 (80.9)	
Tissue IL-5 (pg/ml, mean (SD))		123.63 (244.09)	407.15 (625.97)	<0.01
Tissue IL-17 (pg/ml, mean (SD))		78.99 (112.31)	77.84 (588.03)	0.93
Blood total IgE (KU/L)		235.10 (328.85)	307.85 (588.03)	0.25

Table 1. Baseline comparison of the training set and the validation set. *Abbreviations: EOS count: Blood eosinophil count. NERD: NSAID-Exacerbated Respiratory Disease*

diseases were excluded from the study. Patients' medical records were reviewed for information regarding demographics, clinical characteristics, laboratory test results, and histopathological characteristics. Patient's anosmia is measure by symptom. The characteristics of enrolled patients are shown in Table 1. According to previous studies, eosinophilic CRSwNP (ECRSwNP) was defined by eosinophil infiltration of more than 10/high power field (HPF, 400×).[12] Polyp tissue was snap-frozen in liquid nitrogen and stored at −80 °C for later tissue homogenization. Concentrations of biomarkers including IL-1β, IL-5, IL-6, IL-8, IL-10, IL-17, tumor necrosis factor -α, IFN-γ, transforming growth factor-β, ECP, and total IgE were tested as previously described.[13] Both institutions have performed cluster analysis of the biomarkers and grouped CRSwNP patients into 2 endotypes: "type 2 CRSwNP" and "non-type 2 CRSwNP", which has been previously

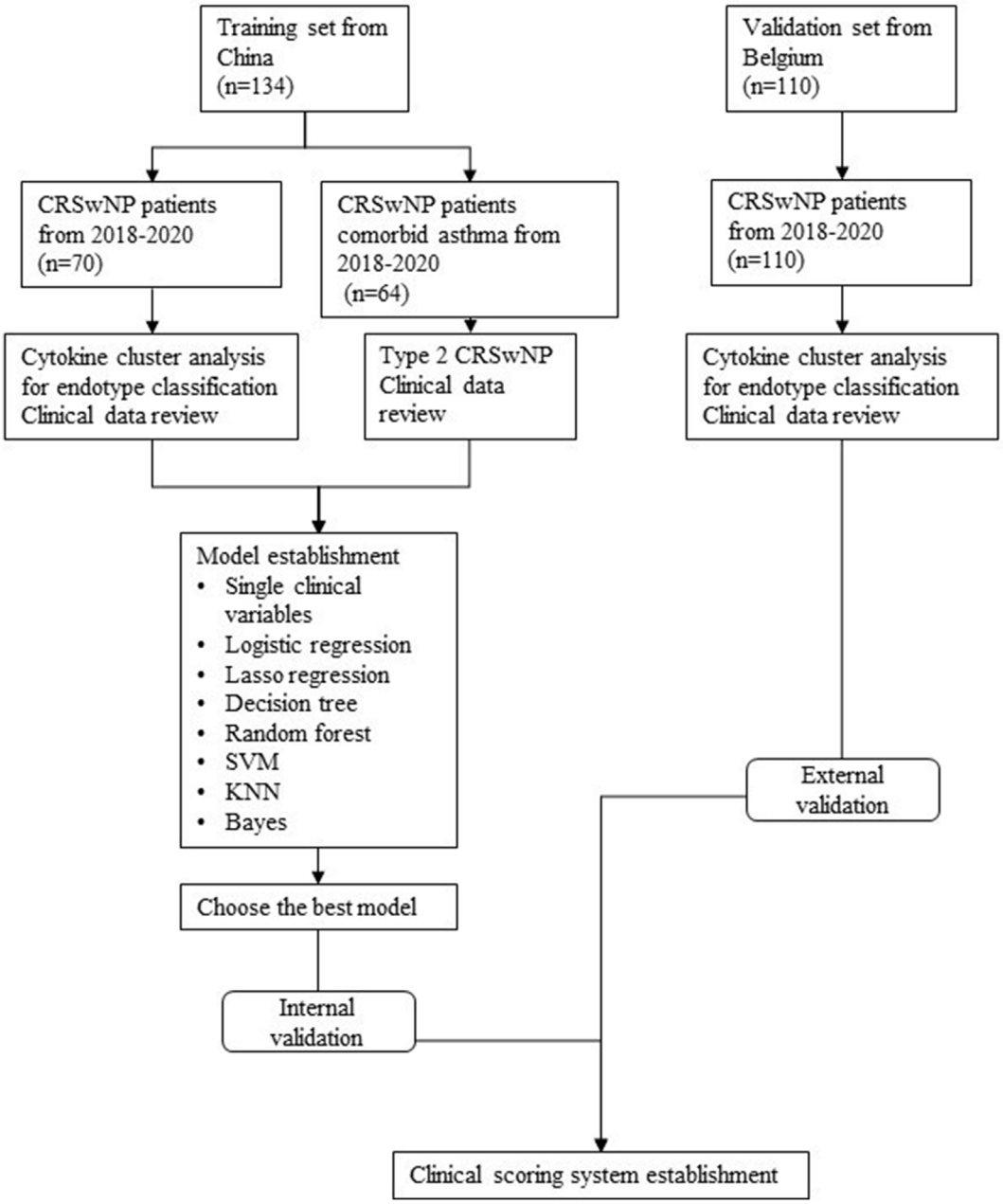

Fig. 1 Flowchart of the study design

demonstrated.[9,13,14] Patients with comorbid asthma were considered as type 2 CRSwNP.[15] Among all the cytokines detected, IL-5 level of polyps distinctly separate patients into 2 groups by setting a threshold of 12.98 pg/g, which was considered as the gold standard for classifying inflammatory endotypes of CRSwNP.[9]

Model development

The flowchart showing the model development and validation process was demonstrated in Fig. 1. We used data from China as the training and internal validation set, and data from Belgium as the external validation set. An IL-5 level of polyp tissue higher than 12.9 pg/g or comorbid asthma was considered as type 2 CRSwNP. Several algorithms using clinical variables were established to diagnose CRSwNP endotype and were later compared to find the best model. First, we evaluate the diagnostic efficacy of single clinical variables including eosinophil (EOS) count, blood eosinophil percentage (EOS percentage), and

ECRSwNP. Second, we adopted a supervised learning approach, which is logistic regression (LR). Variables with $P < 0.1$ in the univariable analysis were selected for multivariate analysis. Stepwise selection in both directions was performed under these criteria. To obtain stabilized coefficient, the training set was resampled by bootstrap algorithm. Third, several machine learning (ML) algorithms, including LASSO regression, random forest (RF), decision tree (DT), support vector machine (SVM), K Nearest Neighbors (KNN), and Naive Bayesian algorithm (Bayes) were used to classify the CRSwNP. The metrics of prediction accuracy and the area under the receiver operating characteristic (ROC) curve (AUC) were used to evaluate model performances. DeLong test was used to compare AUC of each model. We chose the model with the best AUC to establish a clinical scoring system. Clinical variables are converted to scores by the value multiplied by the corresponding weighting coefficient and added to an overall sum.

Statistical analysis

Statistical analysis of the data was performed using R (version 4.1.0) [16]and Rstudio software (version March 1, 1093).[17] Continuous variables with normal distribution were expressed as mean ± standard deviation, and analyzed by Student-t test. Continuous variables with non-normal distribution were expressed as medians and interquartile ranges, and analyzed by the Kruskal-Wallis test. Categorical variables were expressed as frequencies and percentages. Pearson's chi-squared test was used to compare the differences between groups. Differences were considered statistically significant when the P value was less than 0.05. All the R packages used in the model establishment were listed in eTable 1.

RESULTS

Baseline characteristics of the training set and validation set

In total, 134 patients from China as training set and 110 patients from Belgium as validation set were included in this study. The demographic characteristics and clinical information of each group are described in Table 1. Among the 2 datasets, no significant differences were detected regarding age distribution, number of patients with anosmia, mean blood eosinophil count, and number of patients with polyp eosinophil infiltration. In the training set, 57.5% of patients are male, while the majority of patients are male (70.9%) in the validation set. More patients (39.7%) in the validation set had smoking habits. Eighty-nine (66.4%) patients in the training set and 49 (44.5%) in the validation set are comorbid with asthma. There are 110 and 89 type 2 CRSwNP in the training set and validation set, respectively. The training set and validation set are comparable in the baseline characteristics.

Model establishment and evaluation

First, we tested if single clinical variable had the ability to diagnose CRSwNP endotype. The AUC of EOS count and EOS percentage were 76.84% (95%CI = 64.62%–89.06%) and 75.25% (95% CI = 62.79%–87.70%), with a cutoff value of $2.85*10^9$/L and 4%. The AUC of ECRSwNP was 64.55% (95%CI = 53.64%–75.45%).

Next, we tried to adopt LR model or machine learning (ML) model with multiple clinical variables to diagnose CRSwNP endotype. Eight candidate variables were identified, including age, gender, smoking habits, alcohol consumption, symptoms of anosmia, history of CRSwNP recurrence, blood eosinophil count, and polyp eosinophil infiltration levels. Regarding the LR model, univariate analyses demonstrated that age, gender, symptoms of anosmia, blood eosinophil count, and polyp eosinophil infiltration level are likely to be associated with type 2 CRSwNP with $P < 0.1$. We included all of the relevant variables in the multivariate analysis. The detailed results of the univariate and multivariate analysis are presented in eFig. 1. The AUC of the LR model was 83.18%. The ROC threshold for the LR score is 1.2. Machine learning (ML) algorithms including LASSO regression, DT, RF, SVM, KNN, and Naive Bayesian algorithm were applied to establish diagnostic models. The best-performing parameters and the model establishing process were shown in eFigure 2-4. The diagnostic performance, including sensitivity, specificity, AUC, and accuracy of different models were shown in Table 2. The ROC curves of each model were demonstrated in Fig. 2.

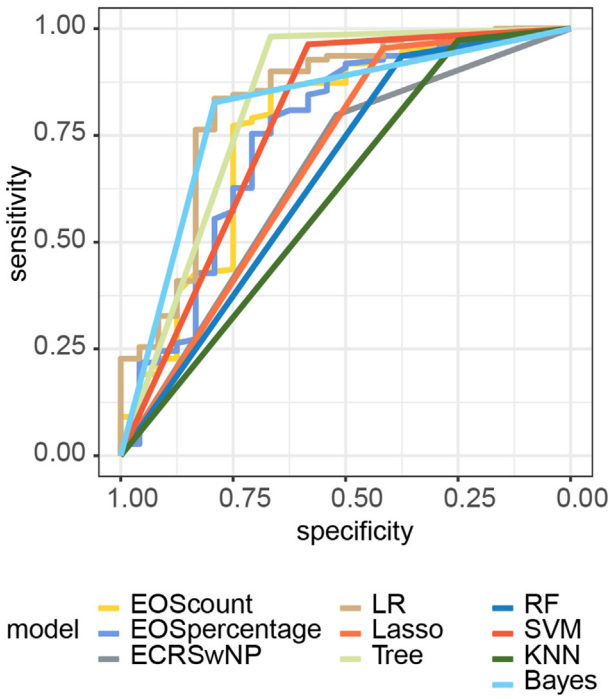

Fig. 2 The ROC curves of all the diagnostic models

Finally, we compared all the models to select the one with best diagnostic performance. The LR model had the highest AUC, and the DT model had the highest accuracy at 90.3%. While comparing the LR model and single clinical variables, both EOS count and EOS percentage showed similar performance to the LR model (LR vs EOS count, P = 0.1751; LR vs EOS percentage, P = 0.1244), but the diagnostic capability of ECRSwNP is worse than LR model (LR vs ECRSwNP, P = 0.0245). The predictive performances of the LR, Bayes, and DT model were similar (AUC = 83.18%, 80.95%, 82.42%, respectively; LR vs Bayes, P = 0.3695; LR vs DT, P = 0.8596) and outperformed the LASSO, RF, SVM, and KNN model (AUC = 68.56%, 65.57%, 77.35%, 61.14%, respectively; LR vs LASSO, P = 0.0458; LR vs RF, P < 0.001; LR vs SVM, P = 0.0201; LR vs KNN, P < 0.001).

Model validation and scoring system development

Although the diagnostic efficacy of the LR model was similar to EOS count and EOS percentage, it included all the relevant clinical characteristics of patients. In addition, the coefficients of an LR formula can be easily extracted, interpreted, and shared. Therefore, we chose LR model for external validation and transform it into a clinical scoring system. In the validation set, the AUC of LR model is 80.10% (95%CI = 69.77%–90.43%), with sensitivity of 77.53%, specificity of 61.90%, and accuracy of 74.55%. The performance of the LR model was internally and externally validated by bootstrapping algorithm. The calibration curves corresponded to the ideal plot (the 45°line), which revealed a favorable consistency between the LR model estimation and actual observation regarding the diagnosis performance (eFig. 5). These results suggested that the LR model was reasonably accurate, and repeatable. To develop a clinical scoring system, the individual score was calculated based on (LR

	AUC	CI	accuracy	CI	sensitivity	specificity
EOS count	0.7684	0.6462–0.8906	0.7687	0.6880–0.8371	0.7727	0.7500
EOS percentage	0.7525	0.6279–0.8770	0.7313	0.6480–0.8042	0.7364	0.7083
ECRSwNP	0.6455	0.5364–0.7545	0.7388	0.6559–0.8108	0.7909	0.5000
Logistic regression	0.8318	0.8031–0.9998	0.7463	0.6639–0.8174	0.7273	0.8333
Lasso regression	0.6856	0.5830–0.7882	0.8582	0.7875–0.9124	0.9545	0.4167
Decision tree	0.8242	0.7271–0.9214	0.9254	0.8670–0.9636	0.9818	0.6667
Random forest	0.6557	0.5541–0.7572	0.8358	0.7620–0.8942	0.9364	0.375
SVM	0.7735	0.6712–0.8757	0.8955	0.8309–0.9417	0.9636	0.5833
KNN	0.6114	0.5216–0.7012	0.8433	0.7705–0.9003	0.9727	0.2500
Naïve Bayes	0.8095	0.7192–0.8997	0.8209	0.7453–0.8817	0.8273	0.7917

Table 2. Comparison of the diagnostic performance of different models. *Abbreviations: EOS count: Blood eosinophil count; EOS percentage: Blood eosinophil percentage; SVM: support vector machine; KNN: K Nearest Neighbors*

Variables	LR score	LR coefficient	Simple Score
Age		0.046	0.1*age
Gender (Male)	1	−1.301	−3
Blood EOS count (*10^9/L)		4.945	10* Blood EOS count
ECRSwNP (≥10/HPF)	1	0.994	2
Anosmia (Yes)	1	1.448	3

Table 3. The transformation of the LR model into the clinical scoring system. *The threshold for diagnosing Type 2 CRSwNP is 8.9. Example: A 60-year-old female patient has anosmia, with blood EOS count of 0.6*10^9/L and nasal polyp infiltration, her simple score will be calculated as following. Simple score = 0.1*60 (age)+0 (gender)+3 (anosmia)+10*0.6 (blood EOS count)+2 (ECRSwNP) = 17. She should be considered as type 2 CRSwNP*

score*LR coefficient*2) and rounded up (Table 3). Total simple score was calculated according to each individual score. The threshold of simple score for diagnosing Type 2 CRSwNP is 8.9. The diagnostic performance of the scoring system (AUC = 83.37%) and LR model were similar (P = 0.6633). To better illustrate how the scoring system works, an example is provided (eFig. 6).

DISCUSSION

In this study, a variety of diagnostic models of CRSwNP endotype were established based on the clinical and histopathological characteristics of CRSwNP patients. To the best of our knowledge, this is the first study to apply cytokine measurement as the gold standard for CRSwNP endotype model establishment and validation in a setting of multinational institutions. Among all the models, LR models had the best diagnostic performance with AUC of 83.18%. After internal and external validation, the LR model presented satisfactory reliability. In addition, to facilitate the clinical usability, the LR model was transformed into a simple scoring system, which showed comparable diagnostic efficacy to the LR model. This user-friendly model can help clinicians classify the endotype of CRSwNP patients, and develop precise and personalized treatment plans.

The CRS endotype plays a significant role in developing treatment strategies and predicting patients' prognoses. According to the latest CRS treatment guidelines, topical and oral corticosteroid treatment is recommended for CRSwNP patients.[18–22] However, this treatment strategy is ineffective in approximately 38%–51% of CRS patients, suggesting the heterogeneity of CRSwNP.[23,24] Most of the severe CRSwNP patients display type 2 characteristics with high recurrence rates, and difficulties in controlling their disease with conventional surgeries and medicine, which has drawn the attention of clinicians and researchers. Accurately distinguishing the endotype can help in individualized treatment decision for CRSwNP patients, and also assist in clinical research on biologic treatments. In line with previous studies, clustering analysis of cytokine levels suggested that the presence of IL-5 is the paramount factor indicating the endotype with CRSwNP and asthma.[9,13,25] A number of large-scale prospective and randomized clinical trials have proved that mepolizumab, a humanized monoclonal antibody against IL-5, has satisfactory therapeutic effects on CRSwNP and asthma.[26–28] Therefore, the practical and accurate identification of type 2 CRSwNP based on cytokine levels allows patients to receive individualized treatment strategies.

Among the clinical features of CRSwNP patients, a simple way to distinguish type 2 inflammation is whether the patient has comorbid late-onset asthma according to European studies, but no similar data is reported in Asian population.[15] However, the order of onset regarding CRSwNP and asthma has not yet been determined. A number of patients with recurrent CRSwNP did not have asthma at their first diagnosis, thus missing the optimal timing for intervention. For the pathological features, eosinophil infiltration in NP is often used to classify CRSwNP into ECRSwNP and non-ECRSwNP.[6,29,30] A large number of studies have shown that patients with ECRSwNP are often comorbid with asthma, low quality of life, and high recurrence rate after

conventional treatments.[22,23,25] However, the diagnostic criteria of ECRSwNP have not yet been unified. The commonly used classification criteria for ECRSwNP are: ①Eosinophils accounted for >25% of total cells in NP;[31] ②The number of eosinophils ≥10/HPF.[32] For polyps with a high degree of edema or with aggregation of eosinophils, percentage of eosinophils would be more precisely reflecting the eosinophilic degree. However, this method requires counting the total number of cells per high-power field of view, which is time-consuming and labor-intensive. Currently, the development of artificial intelligence in the identification of histology slides would help to promote the application of these criteria.[33] The second criteria are relatively simple and popular. Meanwhile, European Position Paper on Rhinitis and Nasal Polyps (EPOS 2020) also supports this method to classify ECRSwNP.[11] However, the second criteria are prone to have observation bias. It has also been suggested that further stratifications of the eosinophil number would help to determine the severity of type 2 inflammation. Unlike histopathological characteristics, cytokine-based endotyping can completely and accurately reflect the predominant immune status of nasal mucosa.[9,20] However, this classification scheme is mainly applied in research settings due to the high expense and invasiveness.

In this study, the diagnostic model integrated demographic information, clinical symptoms, laboratory tests, and histopathological characteristics of CRSwNP patients, which is able to comprehensively reflect patients' inflammatory endotypes. In the LR model, age, gender, anosmia, blood eosinophil count, and tissue eosinophilic infiltration level are chosen as variables. The previous study suggested that elderly patients had a higher risk of developing type 2 CRSwNP.[34] At present, there is no consensus on the effect of gender on the inflammatory endotypes, but animal experiments have found that females are more prone to type 2 inflammation-predominant diseases.[35–37] Polyps arising from olfactory cleft have a greater impact on olfactory function. Therefore, although CT scores were not added to our model for the convenience of clinical application, the symptom of anosmia can imply polyps from olfactory cleft to a certain extent. Corresponding to the coefficient in LR model, blood eosinophil count weights greater than tissue eosinophil infiltration, suggesting that type 2 inflammation in CRSwNP is not limited to the nasal cavity but also has systemic effects. Studies have shown that when peripheral blood eosinophils increase, eosinophils are more likely to be activated with a more vigorous oxidative metabolism process, more easily regulated by IL-5, and highly expressed CD49d, CCR3, and CD25 receptors.[38] Wang et al explored the relationship between systemic and local eosinophilia and found that these 2 characteristics are independent factors for poor disease control of CRSwNP.[39] This result was in line with our study and suggested that type 2 CRSwNP has a worse prognosis.

Our study establishes several diagnostic models based on different algorithms. Through comparison, the diagnostic performance of the LR model is better than others. As the most commonly used model in prediction and diagnosis, LR is able to classify the dataset without assuming data distribution in advance. In addition, the coefficients of an LR formula can be easily extracted, interpreted, and shared. However, it also has the disadvantage of overfitting the collinear variables.[40–42] Lasso regression can improve overfitting problems by adjusting parameters and simplifying the model, which is suitable for datasets with high dimensions and correlation. However, the variables in our study were not enough to establish LASSO models without overfitting. For ML algorithms, each patient would be classified into either type 2 or non-type 2 CRSwNP, and the ROC curves of these models were polyline. We suggested that ML can better identify type 2 CRSwNP, with a higher positive predictive value, while the sensitivity and specificity of the LR model are in a more balanced manner.[43,44] The positive predictive value and negative predictive value of the diagnostic model were related to prevalence of type 2 CRSwNP. The positive predictive value would be higher when the prevalence of type 2 CRSwNP is higher. In this study, we established a clinical scoring system based on the LR model, by converting complex regression formulas into simple equations, which

would facilitate the application of the diagnostic model in clinical practice.

This study has certain limitations. Firstly, this study is a retrospective study, which inevitably had selection bias. In order to accurately measure the cytokine levels in polyp tissue, patients with CRSwNP who applied topical and systemic steroids before sample collection were excluded. Secondly, the study population is relatively small, and the cutoff value needs to be modified in a larger clinical setting. Lastly, we are unable to integrate non-type 2 cytokine profile into the diagnostic criteria, which may have resulted in a minor bias in the inflammatory endotyping process.

In this study, multiple CRSwNP endotype diagnostic models were established based on the cytokine clustering analysis as gold standard. By integrating the clinical and histopathological characteristics of CRSwNP patients, the LR model has the best diagnostic performance and showed good stability after internal and external validation. The clinical scoring system based on LR model can assist clinicians in quickly classifying the inflammatory endotypes of patients with CRSwNP and facilitating the development of individual treatment plans.

Abbreviations

CRSsNP: CRS without nasal polyps; CRSwNP: CRS with nasal polyps; ECRSwNP: eosinophilic CRSwNP; EOS count: blood eosinophil count; EOS percentage: blood eosinophil percentage; LR: logistic regression; ML: machine Learning; RF: random forest; DT: decision tree; SVM: support vector machine; KNN: K Nearest Neighbors; Bayes: Naive Bayesian algorithm; ROC: the receiver operating characteristic; AUC: the area under the receiver operating characteristic curve

Funding source

This work was supported by the National Natural Science Foundation of China (NSFC) grants (81900918, 82020108009, 81870696), Natural Science Foundation of Guangdong Province, China (2022A1515010002. 2018B030312008), Guangdong Research Program of Key Fields in Province (2020B1111300003) and the Key-Area Research and Development of Guangdong Province (2020B1111190001).

Availability of data and material

All data generated or analyzed during this study are included in this published article and its Supplemental file. More related data of the current study are available from the corresponding author on reasonable request.

Author contribution

(I) Conception and design: Wei-Ping Wen, Yi-Hui Wen; Claus Bachert.
(II) Administrative support: Wei-Ping Wen, Yi-Hui Wen; Claus Bachert.
(III) Provision of study materials or patients: Wei-Ping Wen, Yi-Hui Wen, Zhao-Feng Xu.
(IV) Collection and assembly of data: Meng-Yu Chen, Zhao-Feng Xu, Tong Lu, Zheng-Qi Li, Jian Li: Nan Zhang; Claus Bachert.
(V) Data analysis and interpretation: Meng-Yu Chen, Zhao-Feng Xu, Yi-Wei Fu.
(VI) Manuscript writing: All authors.
(VII) Final approval of manuscript: All authors.

Ethics approval

Ethics approved by Ethics Committee for Clinical Research and animal trials of the First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China approved the study (Number: [2020]302). The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by ethics board of the First Affiliated Hospital of Sun Yat-sen University ([2020] 302), and written informed consent was obtained from all subjects before inclusion.

Consent for publication

All authors have seen and approved the last version and agreed to publication of the work.

Declaration of competing interest

All authors have completed the ICMJE uniform disclosure form. Dr. Yi Hui Wen reports that he is funded by the National Natural Science Foundation of China (NSFC) grants (81900918) and Guangdong Natural Science Foundation of China grants (2022A1515010002). Dr. Wei-Ping Wen reported that he was funded by the National Natural Science Foundation of China (NSFC) grants (82020108009, 81870696), Guangdong Natural Science Foundation of China grants (2018B030312008), Guangdong Research Program of Key Fields in Province (2020B1111300003) and the Key-Area Research and Development of Guangdong Province (2020B1111190001).

Acknowledgments

The authors express appreciation for the participation of all people who contributed to the design, analysis, and interpretation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.waojou.2023.100796.

Author details

[a]Department of Otolaryngology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong, PR China. [b]Otorhinolaryngology Institute, Sun Yat-sen

University, Guangzhou, Guangdong, PR China. [c]The First Affiliated Hospital, Sun Yat-sen University, International Airway Research Center, Guangzhou, Guangdong, PR China. [d]Department of Musculoskeletal Oncology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China. [e]The Upper Airways Research Laboratory, Department of Oto-Rhino-Laryngology, Ghent University Hospital, Ghent, Belgium. [f]Department of Otolaryngology, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, PR China. [g]First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, PR China.

REFERENCES

1. Shi Jb, Fu QL, Zhang H, et al. Epidemiology of chronic rhinosinusitis: results from a cross-sectional survey in seven Chinese cities. *Allergy*. 2015;70(5):533–539. https://doi.org/10.1111/all.12577.
2. Hastan D, Fokkens WJ, Bachert C, et al. Chronic rhinosinusitis in Europe – an underestimated disease. A GA2LEN study. *Allergy*. 2011;66(9):1216–1223. https://doi.org/10.1111/j.1398-9995.2011.02646.x.
3. Tomassen P, Newson RB, Hoffmans R, et al. Reliability of EP3OS symptom criteria and nasal endoscopy in the assessment of chronic rhinosinusitis–a GA^2 LEN study. *Allergy*. 2011;66(4):556–561. https://doi.org/10.1111/j.1398-9995.2010.02503.x.
4. Van Zele T, Claeys S, Gevaert P, et al. Differentiation of chronic sinus diseases by measurement of inflammatory mediators. *Allergy*. 2006;61(11):1280–1289. https://doi.org/10.1111/j.1398-9995.2006.01225.x.
5. Wang X, Zhang N, Bo M, et al. Diversity of TH cytokine profiles in patients with chronic rhinosinusitis: a multicenter study in Europe, Asia, and Oceania. *J Allergy Clin Immunol*. 2016;138(5):1344–1353. https://doi.org/10.1016/j.jaci.2016.05.041.
6. Kim SJ, Lee KH, Kim SW, Cho JS, Park YK, Shin SY. Changes in histological features of nasal polyps in a Korean population over a 17-year period. *Otolaryngol Head Neck Surg*. 2013;149(3):431–437. https://doi.org/10.1177/0194599813495363.
7. Wang W, Gao Y, Zhu Z, et al. Changes in the clinical and histological characteristics of Chinese chronic rhinosinusitis with nasal polyps over 11 years. *Int Forum Allergy Rhinol*. 2019;9(2):149–157. https://doi.org/10.1002/alr.22234.
8. Jiang WX, Cao PP, Li ZY, et al. A retrospective study of changes of histopathology of nasal polyps in adult Chinese in central China. *Rhinology*. 2019;57(4):261–267. https://doi.org/10.4193/Rhin18.070.
9. Tomassen P, Vandeplas G, Zele TV, et al. Inflammatory endotypes of chronic rhinosinusitis based on cluster analysis of biomarkers. *J Allergy Clin Immunol*. 2016;137(5):1449–1456.e4. https://doi.org/10.1016/j.jaci.2015.12.1324.
10. Wang ET, Zheng Y, Liu PF, Guo LJ. Eosinophilic chronic rhinosinusitis in east asians. *World J Clin Cases*. 2014;2(12):873–882. https://doi.org/10.12998/wjcc.v2.i12.873.
11. Fokkens WJ, Lund VJ, Hopkins C, et al. European position paper on rhinosinusitis and nasal polyps 2020. *Rhinology*. 2020;58(Suppl S29):1–464. https://doi.org/10.4193/Rhin20.600.
12. Cao PP, Li HB, Wang BF, et al. Distinct immunopathologic characteristics of various types of chronic rhinosinusitis in adult Chinese. *J Allergy Clin Immunol*. 2009;124(3):478–484. https://doi.org/10.1016/j.jaci.2009.05.017, 484.e1-2.
13. Liu W, Liu X, Zhang N, et al. Microvessel quantification by fully convolutional neural networks associated with type 2 inflammation in chronic rhinosinusitis. *Ann Allergy Asthma Immunol*. 2022;128(6):697–704.e1. https://doi.org/10.1016/j.anai.2022.02.025.
14. Förster-Ruhrmann U, Szczepek AJ, Pierchalla G, et al. Chemokine expression-based endotype clustering of chronic rhinosinusitis. *J Personalized Med*. 2022;12(4):646. https://doi.org/10.3390/jpm12040646.
15. Bachert C, Han JK, Wagenmann M, et al. EUFOREA expert board meeting on uncontrolled severe chronic rhinosinusitis with nasal polyps (CRSwNP) and biologics: definitions and management. *J Allergy Clin Immunol*. 2021;147(1):29–36. https://doi.org/10.1016/j.jaci.2020.11.013.
16. R Core Team. *Foundation for Statistical Computing. R: a language and environment for statistical computing*; 2019. https://www.rproject.org/. Accessed February 10, 2019.
17. RStudio Team. *RStudio: Integrated Development Environment for R*; 2015. http://www.rstudio.com/. Accessed February 10, 2019.
18. Wallwork B, Coman W, Mackay-Sim A, Greiff L, Cervin A. A double-blind, randomized, placebo-controlled trial of macrolide in the treatment of chronic rhinosinusitis. *Laryngoscope*. 2006;116(2):189–193. https://doi.org/10.1097/01.mlg.0000191560.53555.08.
19. Cardell LO, Stjärne P, Jonstam K, Bachert C. Endotypes of chronic rhinosinusitis: impact on management. *J Allergy Clin Immunol*. 2020;145(3):752–756. https://doi.org/10.1016/j.jaci.2020.01.019.
20. Bachert C, Zhang N, Hellings PW, Bousquet J. Endotype-driven care pathways in patients with chronic rhinosinusitis. *J Allergy Clin Immunol*. 2018;141(5):1543–1551. https://doi.org/10.1016/j.jaci.2018.03.004.
21. Gurrola J, Borish L. Chronic rhinosinusitis: endotypes, biomarkers, and treatment response. *J Allergy Clin Immunol*. 2017;140(6):1499–1508. https://doi.org/10.1016/j.jaci.2017.10.006.
22. Liao B, Liu JX, Li ZY, et al. Multidimensional endotypes of chronic rhinosinusitis and their association with treatment outcomes. *Allergy*. 2018;73(7):1459–1469. https://doi.org/10.1111/all.13411.
23. Lou H, Meng Y, Piao Y, Wang C, Zhang L, Bachert C. Predictive significance of tissue eosinophilia for nasal polyp recurrence in the Chinese population. *Am J Rhinol Allergy*. 2015;29(5):350–356. https://doi.org/10.2500/ajra.2015.29.4231.
24. Wen W, Liu W, Zhang L, et al. Increased neutrophilia in nasal polyps reduces the response to oral corticosteroid therapy. *J Allergy Clin Immunol*. 2012;129(6):1522–1528.e5. https://doi.org/10.1016/j.jaci.2012.01.079.
25. Turner JH, Chandra RK, Li P, Bonnet K, Schlundt DG. Identification of clinically relevant chronic rhinosinusitis endotypes using cluster analysis of mucus cytokines. *J Allergy Clin Immunol*. 2018;141(5):1895–1897.e7. https://doi.org/10.1016/j.jaci.2018.02.002.
26. Bachert C, Sousa AR, Han JK, et al. Mepolizumab for chronic rhinosinusitis with nasal polyps: treatment efficacy by comorbidity and blood eosinophil count. *J Allergy Clin*

Immunol. 2022;149(5):1711–1721.e6. https://doi.org/10.1016/j.jaci.2021.10.040.

27. Fokkens WJ, Mullol J, Kennedy D, et al. Mepolizumab for chronic rhinosinusitis with nasal polyps (SYNAPSE): in-depth sinus surgery analysis. *Allergy*. July 2022. https://doi.org/10.1111/all.15434.
28. Gevaert P, Van Bruaene N, Cattaert T, et al. Mepolizumab, a humanized anti-IL-5 mAb, as a treatment option for severe nasal polyposis. *J Allergy Clin Immunol*. 2011;128(5):989–995. https://doi.org/10.1016/j.jaci.2011.07.056. e1-8.
29. Tokunaga T, Sakashita M, Haruna T, et al. Novel scoring system and algorithm for classifying chronic rhinosinusitis: the JESREC Study. *Allergy*. 2015;70(8):995–1003. https://doi.org/10.1111/all.12644.
30. Snidvongs K, Lam M, Sacks R, et al. Structured histopathology profiling of chronic rhinosinusitis in routine practice. *Int Forum Allergy Rhinol*. 2012;2(5):376–385. https://doi.org/10.1002/alr.21032.
31. Ma L, Shi J, Wang K, Sun Y, Xu R. Clinical characteristics of patients with CRSwNP with intensely high eosinophil level. *Laryngoscope Investig Otolaryngol*. 2022;7(2):316–324. https://doi.org/10.1002/lio2.758.
32. Liu Z, Chen J, Cheng L, et al. Chinese society of allergy and Chinese society of otorhinolaryngology-head and neck surgery guideline for chronic rhinosinusitis. *Allergy, Asthma & Immunology Research*. 2020;12(2):176–237. https://doi.org/10.4168/aair.2020.12.2.176.
33. Wu Q, Chen J, Ren Y, et al. Artificial intelligence for cellular phenotyping diagnosis of nasal polyps by whole-slide imaging. *EBioMedicine*. 2021;66. https://doi.org/10.1016/j.ebiom.2021.103336.
34. Bachert C, Marple B, Schlosser RJ, et al. Adult chronic rhinosinusitis. *Nat Rev Dis Prim*. 2020;6(1):86. https://doi.org/10.1038/s41572-020-00218-1.
35. Laffont S, Blanquart E, Savignac M, et al. Androgen signaling negatively controls group 2 innate lymphoid cells. *J Exp Med*. 2017;214(6):1581–1592. https://doi.org/10.1084/jem.20161807.
36. Melgert BN, Postma DS, Kuipers I, et al. Female mice are more susceptible to the development of allergic airway inflammation than male mice. *Clin Exp Allergy*. 2005;35(11):1496–1503. https://doi.org/10.1111/j.1365-2222.2005.02362.x.
37. Blacquière MJ, Hylkema MN, Postma DS, Geerlings M, Timens W, Melgert BN. Airway inflammation and remodeling in two mouse models of asthma: comparison of males and females. *Int Arch Allergy Immunol*. 2010;153(2):173–181. https://doi.org/10.1159/000312635.
38. Dupuch V, Tridon A, Ughetto S, et al. Activation state of circulating eosinophils in nasal polyposis. *Int Forum Allergy Rhinol*. 2018;8(5):584–591. https://doi.org/10.1002/alr.22079.
39. Wang K, Deng J, Yang M, et al. Concordant systemic and local eosinophilia relates to poorer disease control in patients with nasal polyps. *World Allergy Organ J*. 2019;12(8), 100052. https://doi.org/10.1016/j.waojou.2019.100052.
40. Cole TJ. Applied logistic regression. D. W. Hosmer and S. Lemeshow, Wiley, New York, 1989. No. of Pages: xiii + 307. Price: £36.00. *Statistic Med*. 1991;10(7):1162–1163. https://doi.org/10.1002/sim.4780100718.
41. Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat Sci*. 1986;1(1):54–75. https://doi.org/10.1214/ss/1177013815.
42. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985;39(4):783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x.
43. Song X, Liu X, Liu F, Wang C. Comparison of machine learning and logistic regression models in predicting acute kidney injury: a systematic review and meta-analysis. *Int J Med Inf*. 2021;151, 104484. https://doi.org/10.1016/j.ijmedinf.2021.104484.
44. Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. *Value Health*. 2019;22(7):808–815. https://doi.org/10.1016/j.jval.2019.02.012.