

Development and multicenter validation of a novel radiomics-based model for identifying eosinophilic chronic rhinosinusitis with nasal polyps*

Ke-Zhang Zhu^{1,2,3,#}, Chao He^{1,2,3,#}, Zhen Li⁴, Peng-Ju Wang⁵, Shu-Xin Wen⁶, Kai-Xue Wen⁶, Jing-Yuan Wang⁶, Jun Liu⁵, Heng Xiao^{1,2,3}, Cui-Lian Guo^{1,2,3}, Ao-Nan Chen^{1,2,3}, Jin-Hua Zhang⁴, Xiang Lu^{1,2,3,†}, Ming Zeng^{1,2,3,†}, Zheng Liu^{1,2,3,†}

Rhinology 61: 2, 132 - 143, 2023
<https://doi.org/10.4193/Rhin22.361>

***Received for publication:**
 September 11, 2022

Accepted: December 21, 2022

¹ Department of Otolaryngology-Head and Neck Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, P.R. China

² Institute of Allergy and Clinical Immunology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, P.R. China

³ Hubei Clinical Research Center for Nasal Inflammatory Diseases, Wuhan, P.R. China

⁴ Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, P.R. China

⁵ Department of Otolaryngology-Head and Neck Surgery, Xiangyang Central Hospital, Affiliated Hospital of Hubei University of Arts and Science, Xiangyang, P.R. China

⁶ Department of Otolaryngology-Head and Neck Surgery, Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Tongji Shanxi Hospital, Third Hospital of Shanxi Medical University, Taiyuan, P.R. China

Shared first authorship

† Joint corresponding authors

Abstract

Background: Reliable noninvasive methods are needed to identify endotypes of chronic rhinosinusitis with nasal polyps (CRSwNP) to facilitate personalized therapy. Previous computed tomography (CT) scoring system has limited and inconsistent performance in identifying eosinophilic CRSwNP. We aimed to develop and validate a radiomics-based model to identify eosinophilic CRSwNP.

Methods: Surgical patients with CRSwNP were recruited from Tongji Hospital and randomly divided into training (n = 232) and internal validation cohort (n = 61). Patients from two additional hospitals served as external validation cohort-1 (n = 84) and cohort-2 (n = 54), respectively. Data were collected from October 2013 to May 2021. Eosinophilic CRSwNP was determined by histological criterion. The least absolute shrinkage and selection operator and the logistic regression (LR) algorithm were used to develop a radiomics model. Univariate and multivariate LR were employed to build models based on CT scores, clinical characteristics, and the combination of radiological and clinical characteristics. Model performance was evaluated by assessing discrimination, calibration, and clinical utility.

Results: The radiomics model based on 10 radiomic features achieved an area under the curve (AUC) of 0.815 in the training cohort, significantly better than the CT score model based on ethmoid-to-maxillary sinus score ratio with an AUC of 0.655. The combination of radiomic features and blood eosinophil count had a further improved performance, achieving an AUC of 0.903. The performance of these models was confirmed in all validation cohorts with satisfying predictive calibration and clinical application value.

Conclusion: A CT radiomics-based model is promising to identify eosinophilic CRSwNP. This radiomics-based method may provide novel insights in solving other clinical concerns, such as guiding personalized treatment and predicting prognosis in patients with CRSwNP.

Key words: chronic rhinosinusitis with nasal polyps, computed tomography, endotype, prediction, radiomics

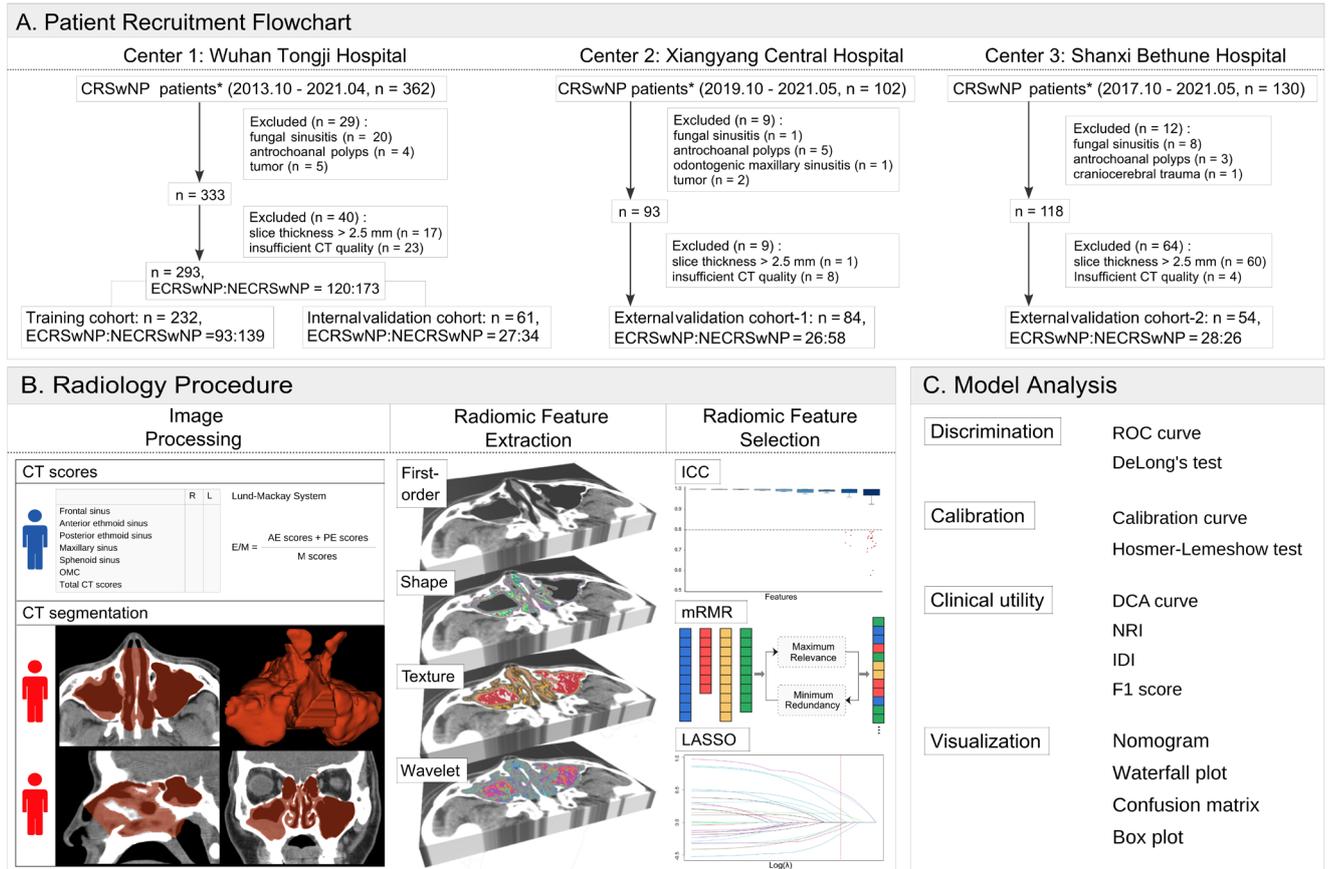


Figure 1. The schematic diagram shows the study processes. A. Patient inclusion and exclusion flowchart, showing four cohorts from three centers with the proportion of eosinophilic and noneosinophilic pathological type. *The inclusion criteria for patient recruitment were: (i) oral glucocorticoid, and intranasal steroid spray and steroid irrigation treatment were discontinued at least 3 months and 1 month before surgery, respectively; (ii) without acute asthma episode or acute upper airway infection within 1 month before surgery; (iii) not under allergen immunotherapy or biologic treatment; B. Radiology flowchart, with image acquisition, image preprocessing (CT score and CT segmentation), radiomic feature extraction, and radiomic feature selection; C. Construction and evaluation of models, including waterfall plot of radiomics model, confusion matrix with box plot of combined model, and ROC, nomogram, calibration curve and DCA of pooled cohorts.

Introduction

Chronic rhinosinusitis (CRS) with nasal polyps (CRSwNP) is a heterogeneous disorder of inflammation of the sinonasal mucosa, which imposes significant health and socioeconomic burdens worldwide (1,2). Based on the severity of tissue eosinophil infiltration, CRSwNP is divided into two subtypes: eosinophilic and noneosinophilic CRSwNP (3,4). These 2 subtypes differ significantly in immunopathological and clinical features and demonstrate distinct responses to glucocorticoid and macrolide treatment, endoscopic sinus surgery, and potential biologics targeting type 2 cytokines (1,2,5,6). Therefore, it is crucial to distinguish eosinophilic and noneosinophilic CRSwNP, for the purpose to optimize the treatment and avoid unnecessary health expenditure.

The gold standard of diagnosis of eosinophilic CRSwNP is the quantification of tissue eosinophil by histopathological methods. However, this method requires invasive tissue sampling

and is time-consuming. Thus, several attempts have been made to identify noninvasive markers or models for diagnosing eosinophilic CRSwNP. Asthma comorbidity (2,6,7), Lund-Mackay computed tomography (CT) score (7,8), blood eosinophil count and percentage (7,9), nasal secretion Charcot-Leyden crystal (CLC) levels (10), and the Japanese Epidemiological Survey of Refractory Eosinophilic Chronic Rhinosinusitis (JESREC) scoring system have been found to be valuable in identifying eosinophilic CRSwNP (11), but with limited or inconsistent performance. In addition, previous studies on these markers have poor quality, particularly lacking internal and external validation. Currently, CT scan is the gold standard in the radiologic evaluation of CRS and conducted under standard procedures in hospitals. The Lund-Mackay score is the most used and validated scoring system of sinonasal inflammatory changes. However, the semi-quantitative nature and narrow scale range (0, 1, and 2) of the Lund-Mackay system make it unable to subgrade the volume of

inflammation and insensitive to change, which is important for characterizing CRS.

Radiomics, a novel machine learning quantitative image analysis method, involves the extraction of high-throughput quantitative features from digitized medical images and has the potential to reveal disease characteristics that cannot be recognized by naked eyes^(12,13). Thus, it may overcome the disadvantages of Lund–Mackay CT scoring. The radiomics has demonstrated promising value in predicting prognosis, gene mutations, and pathological changes in malignant tumors and inflammatory diseases, e.g. differentiating epidermal growth factor receptor mutation status in glioblastoma⁽¹⁴⁾, characterizing intestinal fibrosis in Crohn's disease⁽¹⁵⁾, and predicting the prognosis in interstitial lung disease related to systemic sclerosis⁽¹⁶⁾. To date, there is no attempt to use radiomics in inflammatory upper airway diseases yet. In this study, we aimed to assess the ability of CT radiomics with or without the combination of clinical features to distinguish eosinophilic from noneosinophilic CRSwNP in Chinese patients with multicenter validation.

Methods

Patients and clinical features

This diagnostic study of anonymous data was approved by the Ethics Review Committee at each participating institution. CRSwNP was diagnosed based on the European guideline⁽¹⁾. All the patients had ongoing symptoms after initial attempts on medical treatments and underwent endoscopic sinus surgery. The details of patients including the eligibility criteria are listed in Supplemental Methods. As shown in Figure 1A and Table S1, a total of 431 patients were finally analyzed, with 293 from Center 1 (Tongji Hospital), 84 from Center 2 (Xiangyang Central Hospital), and 54 from Center 3 (Shanxi Bethune Hospital). Patients from Center 1 were randomly divided into training cohort ($n = 232$) and internal validation cohort ($n = 61$) at a ratio of 8:2, while patients from Center 2 and Center 3 served as external validation cohort-1 and external validation cohort-2, respectively. The data were collected from October 2013 to May 2021. The demographic and clinical data including age, sex, comorbidities, previous sinus surgery history, and blood routine test results were retrieved from the medical records anonymously. The symptom score was not included in this study because of its subjectivity. Nasal polyp samples were taken during surgery and subject to routine hematoxylin and eosin staining, and eosinophilic CRSwNP was diagnosed when the percentage of tissue eosinophils exceeded 10% of total infiltrating cells as previously described⁽³⁾.

Radiology procedure

The radiology procedure is shown in Figure 1B. Sinus CT scanning was performed 1–3 days before surgery, and the image

acquisition parameters are shown in Table S2. Lund–Mackay CT scores were determined⁽¹⁾, and the ethmoid-to-maxillary sinus score (E/M) ratio was calculated as:

$$\frac{(\text{anterior ethmoid sinus score} + \text{posterior ethmoid sinus score})}{\text{maxillary sinus score}}^{(17)}$$

CT segmentation was performed on the preprocessed images, and the region of interest (ROI) of an entire sinonasal cavity was semi-automatically delineated using 3D Slicer software (Harvard Medical School, version 4.11.0, <http://www.slicer.org>)⁽¹⁸⁾. Radiomic features (Table S3) were extracted from each ROI using the PyRadiomics package (version 3.0, <https://pyradiomics.readthedocs.io/en/latest>)⁽¹⁸⁾, and selected sequentially using inter-/intra-class correlation coefficients (ICCs), and minimum redundancy maximum relevance (mRMR)⁽¹⁹⁾. Details are shown in Supplementary Methods.

Model construction

For the radiomics models, least absolute shrinkage, and selection operator-logistic regression (LASSO-LR) machine learning algorithms were used for model construction. LR analysis was employed to build models based on the Lund–Mackay CT score, clinical characteristics, and the combination of radiomics and clinical characteristic. Details are provided in Supplementary Methods.

Model analysis

Details are provided in Supplementary Methods and Figure 1C. The performance indicators and assessments of the models mainly include: (i) discrimination: receiver operating characteristic (ROC) curve with the corresponding area under curve (AUC) value and DeLong's test⁽²⁰⁾; (ii) calibration: calibration curve and Hosmer–Lemeshow test⁽²¹⁾; (iii) clinical utility: decision curve analysis (DCA)⁽²²⁾, F1 score, net reclassification improvement (NRI) values, and integrated discrimination improvement (IDI) values^(23,24). The accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), nomogram, waterfall plots, confusion matrices and box plots were also performed.

Sample size consideration

The minimum sample size of training and validation cohorts were calculated according to previous researches^(25,26). Detailed information is provided in the Supplementary Methods.

Statistical analysis

All statistical analysis were performed using Python software (The Python Software Foundation, version 3.7.4) and R software (R Foundation for Statistical Computing, version 4.2.1) (Table S4). All statistical tests were 2-tailed, and $P < 0.05$ was considered statistically significant. More information is provided in the Supplementary Methods.

Table 2. Clinical characteristics of eosinophilic and noneosinophilic CRSwNP patients.

Characteristic	Training cohort (n = 232)		Internal validation cohort (n = 61)		External validation cohort-1 (n = 84)		External validation cohort-2 (n = 54)	
	Eosino- philic CRSwNP (n = 93)	Non eosinophil- ic CRSwNP (n = 139)	Eosino- philic CRSwNP (n = 27)	Non eosinophil- ic CRSwNP (n = 34)	Eosino- philic CRSwNP (n = 26)	Non eosinophil- ic CRSwNP (n = 58)	Eosino- philic CRSwNP (n = 28)	Non eosinophil- ic CRSwNP (n = 26)
Blood routine test								
Leukocyte absolute count, median (IQR), ×10 ⁹ /L	6.59 (5.47-7.95)	6.14 (5.15-7.28)	6.14 (5.32-7.87)	6.45 (5.24-7.48)	6.05 (5.22-7.24)	5.61 (4.85-6.78)	7.15 (5.33-8.23)	5.35 (4.68-7.75)
Neutrophil absolute count, median (IQR), ×10 ⁹ /L	3.36 (2.72-4.24)	3.26 (2.68-4.06)	3.16 (2.36-4.21)	3.35 (2.70-4.11)	3.26 (2.75-3.80)	3.04 (2.49-4.19)	4.00 (2.62-4.37)	3.03 (2.47-4.14)
Neutrophil percent, median (IQR), %	51.8 (46.8-58.9)	54.0 (49.1-60.0)	51.5 (42.8-59.1)	57.3 (50.7-61.8)	52.8 (49.2-57.6)	56.1 (48.6-62.7)	52.8 (47.6-57.6)	53.9 (50.6-58.2)
Lymphocyte absolute count, median (IQR), ×10 ⁹ /L	2.03 (1.65-2.46)	2.10 (1.69-2.53)	2.23 (1.85-2.47)	2.25 (1.59-2.57)	2.22 (1.76-2.72)	1.98 (1.53-2.51)	2.06 (1.59-2.66)	1.89 (1.44-2.36)
Lymphocyte percent, median (IQR), %	32.0 (26.6-38.8)	34.0 (29.6-40.6)	32.7 (28.7-39.1)	31.3 (28.4-35.8)	35.6 (33.1-38.3)	33.6 (28.0-39.9)	31.1 (26.7-37.2)	34.3 (30.4-37.5)
Eosinophil absolute count, median (IQR), ×10 ⁹ /L	0.36 (0.22-0.55)	0.13 (0.08-0.21)	0.37 (0.18-0.66)	0.12 (0.08-0.22)	0.34 (0.19-0.42)	0.11 (0.08-0.22)	0.47 (0.37-0.64)	0.23 (0.05-0.40)
Eosinophil percent, median (IQR), %	5.9 (3.0-8.9)	2.4 (1.4-3.8)	6.6 (2.9-10.6)	2.2 (1.2-3.7)	5.3 (2.5-7.3)	2.3 (1.5-3.3)	7.7 (6.0-9.4)	4.2 (1.1-6.8)
Monocyte absolute count, median (IQR), ×10 ⁹ /L	0.45 (0.36-0.56)	0.45 (0.37-0.57)	0.49 (0.31-0.61)	0.46 (0.35-0.66)	0.37 (0.30-0.46)	0.35 (0.30-0.42)	0.48 (0.36-0.56)	0.37 (0.26-0.58)
Monocyte percent, median (IQR), %	7.2 (6.2-8.4)	7.4 (6.4-9.2)	6.8 (5.8-8.6)	7.6 (6.3-8.6)	5.6 (5.2-6.7)	6.2 (5.2-7.6)	6.4 (5.3-7.8)	7.2 (5.1-7.9)
CT Score, median (IQR)								
Frontal sinus	2 (0-3)	2 (0-3)	1 (0-3)	2 (0-4)	2 (0-3.25)	2 (0.75-4)	2 (0-3)	2 (0-3.25)
Anterior ethmoid sinus	2 (2-4)	2 (2-3)	2 (2-3)	2 (2-3)	2 (2-3)	2 (2-3)	2 (2-3)	2 (1-3.25)
Posterior ethmoid sinus	2 (2-3)	2 (1-3)	2 (2-2)	2 (1-2)	2 (2-2)	2 (1-2.25)	2 (2-3)	2 (0.75-3)
Maxillary sinus	2 (2-3)	2 (2-3)	2 (2-2)	2 (2-3)	2 (2-2)	2 (2-3)	2 (2-2.75)	2 (1.75-2.25)
Sphenoid sinus	1 (0-2)	0 (0-2)	1 (0-2)	0.5 (0-2)	1.5 (0-2)	1 (0-2)	1 (0-2.75)	0.5 (0-2)
OMC	4 (2-4)	4 (2-4)	4 (2-4)	4 (2-4)	4 (4-4)	4 (2-4)	4 (4-4)	4 (2-4)
Total CT score	13 (9-17)	13 (7-17)	12 (7-17)	13.5 (7.5-17.3)	13.5 (10-16.25)	13.5 (9-18)	13 (10.25-17.75)	11 (5.75-16.25)
E/M	2 (2-2.83)	2 (1.3-2.3)	2 (2-2.5)	2 (1.5-2)	2 (2-2)	2 (1.875-2)	2 (2-3)	1.86 (1-2)

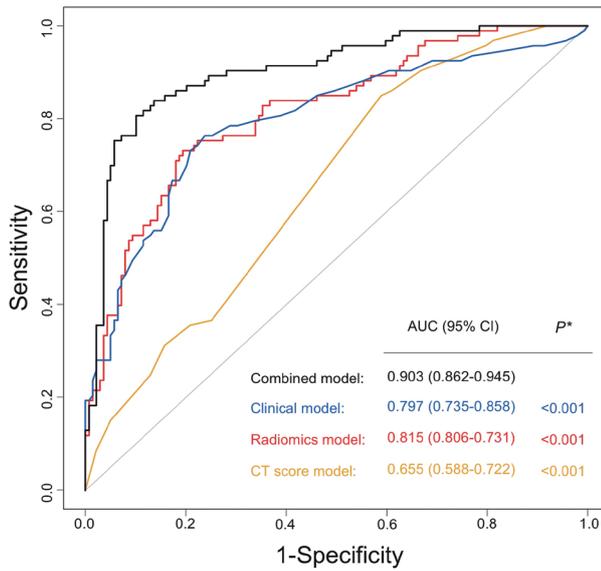
Data are presented as median and interquartile ranges for continuous variables and numbers with percentage for categorical variables and were analyzed by Mann-Whitney U-test. Abbreviations: IQR, interquartile range; OMC, ostiomeatal complex; E/M, (anterior ethmoid sinus score + posterior ethmoid sinus score)/(maxillary sinus score). *The P values in red color indicate those less than 0.05.

Table 3. Predictive capacity of the CT score, radiomics, clinical, and combined model.

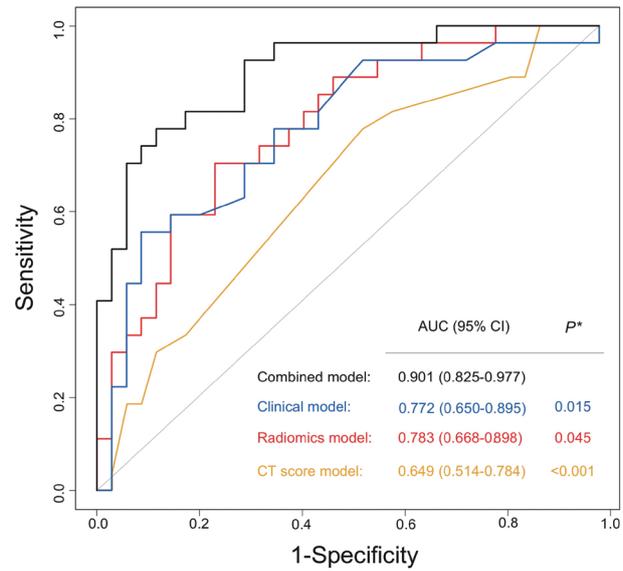
Cohort	Model	AUC (95% CI)	TP	FN	FP	TN	ACC	Sensitivity	Specificity	PPV	NPV	F1 score
Training cohort	CT score model	0.655 (0.588-0.722)	80	13	85	54	0.578	0.860	0.388	0.485	0.806	0.620
	Radiomics model	0.815 (0.759-0.871)	68	25	27	112	0.776	0.731	0.806	0.715	0.817	0.723
	Clinical model	0.797 (0.735-0.858)	71	22	33	106	0.763	0.763	0.763	0.683	0.828	0.721
	Combined model	0.903 (0.862-0.945)	75	18	14	125	0.862	0.806	0.899	0.843	0.874	0.824
Internal validation cohort	CT score model	0.649 (0.514-0.784)	21	6	18	16	0.607	0.778	0.471	0.538	0.727	0.636
	Radiomics model	0.783 (0.668-0.898)	13	14	5	29	0.689	0.481	0.853	0.722	0.674	0.578
	Clinical model	0.772 (0.650-0.895)	17	10	10	24	0.672	0.630	0.706	0.630	0.706	0.630
	Combined model	0.901 (0.825-0.977)	20	7	4	30	0.820	0.741	0.882	0.833	0.811	0.784
External validation cohort-1	CT score model	0.640 (0.524-0.755)	21	5	29	29	0.595	0.808	0.500	0.420	0.853	0.553
	Radiomics model	0.800 (0.699-0.900)	22	4	23	35	0.679	0.846	0.603	0.489	0.897	0.620
	Clinical model	0.771 (0.645-0.896)	19	7	16	42	0.726	0.731	0.724	0.543	0.857	0.623
	Combined model	0.900 (0.830-0.970)	22	4	17	41	0.750	0.846	0.707	0.564	0.911	0.677
External validation cohort-2	CT score model	0.705 (0.572-0.838)	22	6	13	13	0.648	0.786	0.500	0.629	0.684	0.698
	Radiomics model	0.783 (0.661-0.905)	18	10	5	21	0.722	0.643	0.808	0.783	0.677	0.706
	Clinical model	0.771 (0.641-0.901)	25	3	13	13	0.704	0.893	0.500	0.658	0.813	0.758
	Combined model	0.901 (0.820-0.982)	25	3	6	20	0.833	0.893	0.769	0.806	0.870	0.847

Abbreviations: AUC, area under curve; CI, confidence interval; TP, true positive; FP, false positive; FN, false negative; TN, true negative; ACC, accuracy; PPV, positive predictive value; NPV, negative predictive value.

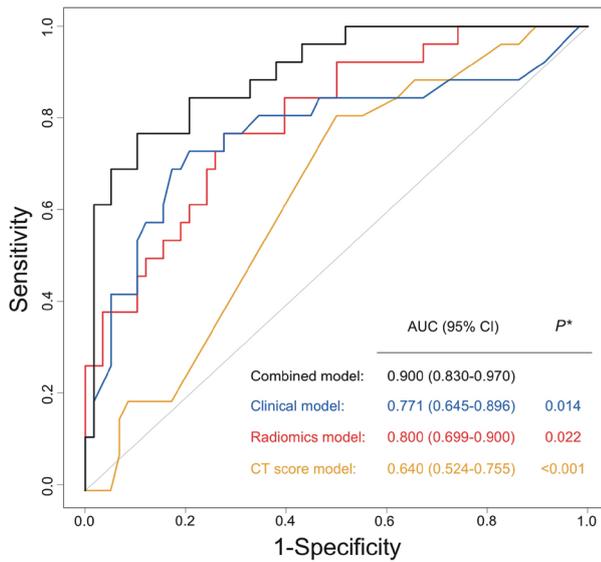
A. Training cohort



B. Internal validation cohort



C. External validation cohort-1



D. External validation cohort-2

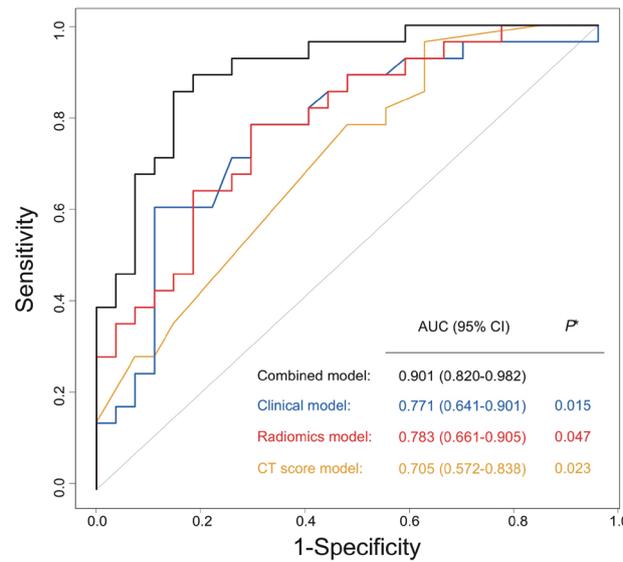


Figure 2. Receiver operating characteristic curves of different models in different cohorts. The values of the areas under the curves were presented in the figures.

(Table 3). The waterfall plot in Figure E2 shows the distribution of rad-scores and endotypes of CRSwNP in the four cohorts, reflecting the good discrimination performance of the LASSO-LR radiomics model. Moreover, eosinophilic and noneosinophilic CRSwNP demonstrated significant differences in the rad-scores in pooled patients with (0.129 [-0.506 to 0.523] vs. -0.652 [-1.206 to -0.236]) and without (0.002 [-0.466 to 0.446] vs. -0.715 [-1.195 to -0.331]) prior surgery (Table S5), indicating a stable distinguishing ability of rad-scores independent of prior surgery.

Clinical model

The demographic and clinical data including age, sex, comorbidities, previous sinus surgery history, and blood routine test results were included in between-group comparison analysis, and asthma comorbidity, peripheral eosinophil absolute count, and eosinophil percentage were found associated with eosinophilic CRSwNP in the training cohort (Table 1 and 2), which was consistent with univariate LR analysis (Table S6). However, only peripheral eosinophil absolute count independently correlated with eosinophilic CRSwNP after backward stepwise multivariate

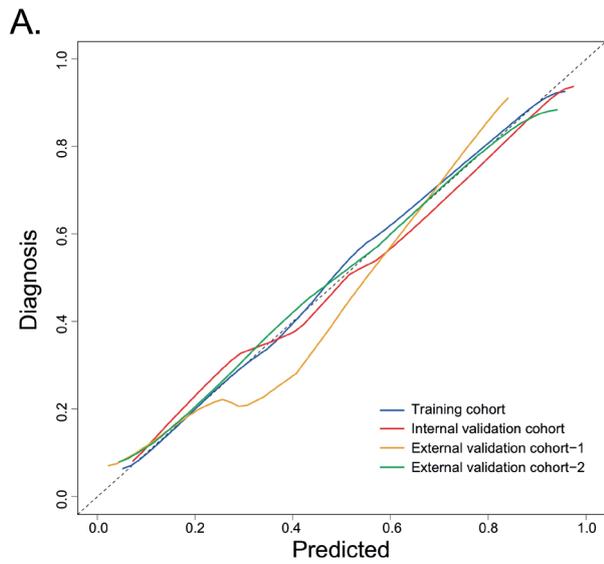
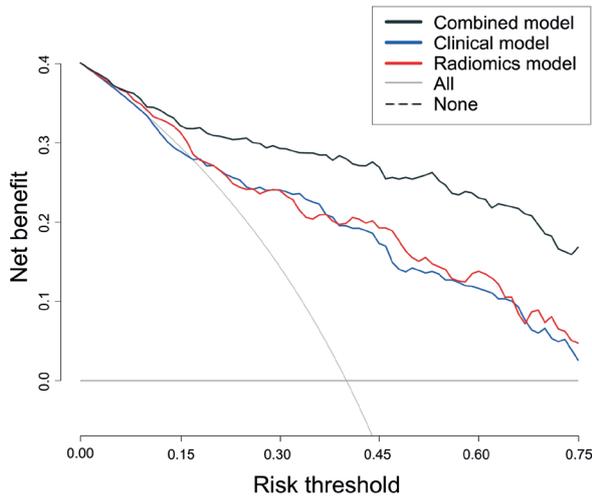
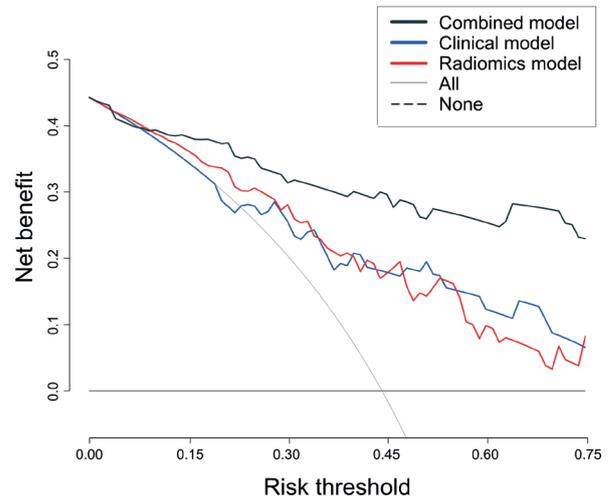


Figure 3. The performance of the combined model. A. Calibration curves of the combined model in different cohorts. The diagonal gray dashed line indicates a perfect prediction, and the solid lines indicate the performance of the models. When the solid line is close to the dotted line, the model works well. B-E. The decision curve analysis of different models in different cohorts. The gray line represents the assumption that “all patients are eosinophilic chronic rhinosinusitis with nasal polyps”, and the dashed black line represents the assumption that “none are eosinophilic chronic rhinosinusitis with nasal polyps”. The combined model shows the highest net benefit compared with the other models in a larger range of threshold probability, indicating the best clinical utility ability.

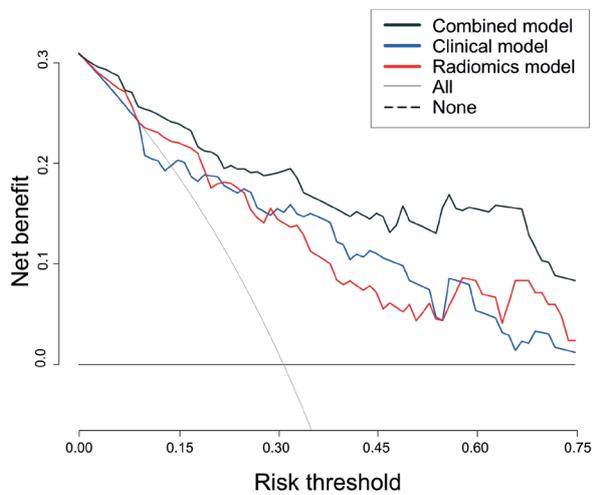
B. Training cohort



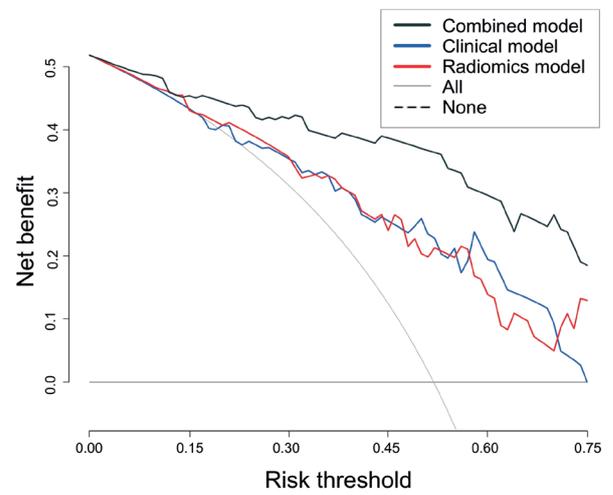
C. Internal validation cohort



D. External validation cohort-1



E. External validation cohort-2



LR analysis (Table S6). A clinical model based on the peripheral eosinophil absolute count provided an AUC of 0.797 (95% CI, 0.735-0.858) in the training cohort, and its moderate diagnostic capacity was validated in internal and external validation cohorts with AUCs ranging from 0.771 to 0.772 (Table 3). The cutoff value was determined to be $0.215 \times 10^9/L$ in the training cohort. There was no significant difference in AUCs between radiomics and clinical model in all cohorts as analyzed by DeLong's test (all $P > 0.05$).

Combined model

We then explored whether integrating radiological and clinical features could achieve better distinguishing efficacy. When integrating CT scores and clinical features, the CT scores features were removed after backward stepwise multivariate LR analysis. Therefore, we focused on the combination of radiomic and clinical features in subsequent analyses.

The correlation analyses showed that there was no or poor correlations between clinical and radiomic features in the training cohort (Figure S3), suggesting that clinical and radiomic features are largely independent of each other and capture the different dimensional characteristics of CRSwNP. The combined model was constructed by integrating rad-score and peripheral eosinophil absolute count using multivariate LR.

$$\text{Risk - score} = 2.140 \times \text{rad-score} + 6.714 \times \text{peripheral eosinophil absolute counts} - 1.387$$

The combined model provided a fair discrimination performance with an AUC > 0.900 in not only the training cohort but also all validation cohorts (Table 3), displaying significant or marginally significant improvement over the radiomics and clinical model (DeLong's test, $P < 0.05$) (Figure 2). The cutoff value of the risk-score in the combined model was determined to be 0.440 based on the Youden's index in the training cohort. Furthermore, the prediction effect of the combined model was evaluated by the confusion matrix and box plots (Figure S4), confirming the fair performance for distinguishing eosinophilic from noneosinophilic CRSwNP in all the cohorts. The combined model achieved comparable efficiency in pooled patients with (AUC, 0.907 [95% CI, 0.844-0.970]) and without (AUC, 0.896 [95% CI, 0.861-0.930]) prior surgery.

Nomogram and calibration curve

To make the combined model more concise and user-friendly, a nomogram was developed based on the rad-score and peripheral eosinophil absolute count with coefficients determined by the multivariate LR analysis (Figure S5). The Hosmer-Lemeshow test showed good calibration of the developed nomogram in all cohorts ($P > 0.05$). The corresponding calibration curves were close to the diagonal dashed line (Figure 3A), indicating a good

agreement between predictive and histologic diagnosis in all cohorts.

Clinical utility

Comparing the combined model with the radiomics or clinical models, the NRI > 0.5 and IDI > 0.1 in all four cohorts indicated a superior reclassification ability of the combined model (Table S7), and the F1 scores of the combined model were also superior over to other models in all cohorts (Table 3).

To further demonstrate the net benefit of predictive models especially the impact of true positive, the DCA was used to compare the benefit of the combined model, radiomics model and clinical model in each cohort (Figure 3B-E). This analysis showed that all models added net benefit than either the "treat all patients as eosinophilic CRSwNP" or "treat none patients as eosinophilic CRSwNP" strategy scheme; however, the combined model significantly increased the net benefit compared to either radiomics model or clinical models over a large range of threshold probability (Figure 3B-E).

Discussion

Eosinophilic CRSwNP is characterized by predominant type 2 inflammation and is associated with a poor prognosis after endoscopic sinus surgery^(5,27), which usually needs more aggressive surgery and postoperative medical treatments. To the best of our knowledge, this is the first study to construct a radiomics-based model for identifying eosinophilic CRSwNP with multicenter validation. The CT radiomics model was significantly superior to conventional CT score model to identify eosinophilic CRSwNP, and further combination with the peripheral blood eosinophil absolute count achieves a fair prediction performance. Our study confers novel insights in a radiomics-based approach in endotype identification and personalized treatment of CRSwNP.

CT is the most used modality for CRS diagnosis and conducted under standard procedure clinically. The Lund-Mackay CT scoring system is widely used for grading CRS severity. Some studies reported that the E/M ratio was an effective predictor of eosinophilic CRSwNP⁽⁸⁾, but others not^(7,28,29). The reasons for those controversial findings may be the subjective nature of image interpretation, and the limited sensitivity of Lund-Mackay scoring system due to its semi-quantitative nature and narrow scale range. In this study, we found that the predictive capacity of E/M ratio for eosinophilic CRSwNP was insufficient (Table 2), confirming that the Lund-Mackay scoring system may be too simple to capture the heterogeneity of CRSwNP. Radiomics, a novel image-based quantitative method, has the potential to overcome the limitations of conventional CT scoring system. Radiomics has been reported to have improved ability to discover the subtle disease heterogeneity, for example, to distinguish the composition of different immune cells infiltra-

ted in advanced solid malignant tumors⁽³⁰⁾. The heterogeneity of inflammatory cell profiles in CRSwNP may also lead to the differences in image intensity, shape, and texture, which can be captured by radiomics. In this study, for the first time, we defined the ROI of an entire sinonasal cavity to obtain as much information of the disease as possible. We developed a radiomics model consisting of 10 key radiomic features to identify eosinophilic CRSwNP by using LASSO-LR algorithm. The AUC of the radiomics model was 0.815 (95% CI: 0.759-0.871) in the training cohort, which was confirmed in all validation cohorts (AUCs, ranging from 0.783 to 0.800), and significantly improved compared to the CT score model. It indicates that the radiomics is superior to conventional Lund–Mackay CT score in capturing disease features. We found these 10 radiomic features can be grouped into 3 categories and provide potentially interpretable information: A) first-order features (minimum, maximum, mean, median and skewness), which represent the gray values of ROI and thus may indicate opacification; B) shape feature (flatness), which represents the shape of ROI; C) texture features (IMC1 and IMC2), which represent the quantification of texture complexity and may reflect the composition of mucosal inflammation such as edema and immune cell infiltration. In order to avoid the bias caused by prior surgery, we excluded bone structure during threshold segmentation and demonstrated significant difference in the rad-scores between eosinophilic and noneosinophilic CRSwNP in both patients with and without prior surgery. Therefore, the validity of our CT radiomics model is unlikely significantly influenced by prior surgery.

In this study, although the radiomics model yielded an enhanced performance compared to conventional Lund–Mackay CT score, its effectiveness was moderate (AUC < 0.850). Our findings are consistent with previous studies employing radiomics to discriminating the pathological features of diseases. Sun et al. reported that the ability of the radiomic signatures to classify high versus low abundance of CD8 cell infiltration in advanced solid tumors was moderate with an AUC of 0.74⁽³⁰⁾. Braman et al. developed a radiomics signature from intra-tumoral region to stratify the human epidermal growth factor receptor 2-enriched and nonenriched breast cancer and achieved a mean AUC of 0.76⁽³¹⁾. Collectively, the findings of our study and other studies indicate that although radiomics has improved ability to capture the complexity of diseases compared to conventional radiological evaluation, it may be unable to reflect all the aspects of the complex diseases. It is therefore necessary to incorporate additional information of other dimensions of the diseases to further improve the performance of the radiomics-based model. Previously, some clinical characteristics including asthma comorbidity, peripheral blood eosinophil count, and loss of smell symptom score have been discovered associated with eosinophilic CRSwNP^(2,6-9,32,33). In this study, to avoid the bias introduced by highly subjective parameters, we decided not to include the

symptom score in our model. Consistent with previous studies^(7,9,33), peripheral eosinophil absolute count was identified as an independent predictor for eosinophilic CRSwNP in our study. The clinical model achieved a performance comparable to radiomics model, but with an AUC value still below 0.85. Notably, we discovered no or poor correlations between radiomic features and clinical characteristics, suggesting that radiomics and clinical features reflect different traits of CRSwNP. We therefore constructed a combined model comprising radiomic features and peripheral eosinophil absolute count and found that the combined model achieved a robust performance with AUCs ranging from 0.900 to 0.903 in all cohorts. By comparing the ROC, DCA, F1 score, NRI and IDI, the performance of the combined model was found superior to the solo radiomics or clinical model. The combined model only requires sinus CT scan and blood routine tests, which are standardized, easily accessible, and exert no extra burden or cost to the patients. We are trying to build a cloud platform which will allow physicians to upload sinus CT images and complete the segmentation of CT images through an internal loop plugin. The software can then automatically complete the radiomics feature extraction and calculation of the risk-score to predict the probability of eosinophilic CRSwNP. Thus, hopefully, our model can provide a noninvasive, reliable, simplified, and clinically accessible tool to identify eosinophilic CRSwNP with high efficacy.

Some limitations have to be addressed in this study. First, the enrolled CRSwNP patients in this study were those refractories to medical treatment and underwent sinus surgery, which cannot represent the whole population of CRSwNP. Nevertheless, these patients are the subgroup most deserving of tailored treatments. Second, although the combined model has been validated in multicenter cohorts, further prospective studies with larger sample sizes are still needed to improve the accuracy and robustness of the model. Third, some characteristics vary between CRSwNP patients in Asian and Western countries. Eosinophilic CRSwNP and asthma comorbidities are more common in Western countries⁽³⁴⁾, so the extrapolation of our model to other populations should be validated with further investigations. Nevertheless, in our study, the external validation cohort-2 had higher atopy and AR comorbidity frequency and peripheral blood eosinophil number (Table S1), suggesting an increased eosinophilic inflammation in this cohort. The validation of the radiomics-based model in this cohort suggests the potential value of our model in a more eosinophilic patient population. Fourth, the segmentation method used in this radiomics study remains somewhat time-consuming. Our future goal is to achieve complete automatic segmentation of ROIs using deep learning algorithms, which may further reduce the time required by the model and facilitate its uptake. However, many samples and extensive calculations are required to achieve this goal.

Conclusions

The radiomics model was significantly superior to the conventional CT score model to identify eosinophilic CRSwNP. The combination of radiomic features and blood eosinophil count has an improved performance over the solo radiomics and clinical models, achieving satisfying performance. This radiomics-based method may also provide novel insights in solving other clinical concerns, such as guiding personalized treatment and predicting prognosis in patients with CRSwNP.

Acknowledgements

None

Authorship contribution

KZZ and CH collected and analyzed data, and prepared the manuscript; PJW, KXW, SXW, JYW and JL collected data for external validation; ZL and JHZ completed image processing and

participated in data discussion; HX, CLG and ANC participated in data retrieval; XL, MZ, and ZL designed the study, interpreted data, and prepared the manuscript.

Conflicts of interest

The authors have no competing interests to declare.

Funding

This study was supported by the Key Research and Development Program of Hubei Province 2021BCA119 (Z.L.), the Natural Science Foundation of Hubei Province grant 2021CFB413 (X.L.), and the National Natural Science Foundation of China (NSFC) grants 82071025 (M.Z.), and 82130030 and 8192010801 (Z.L.).

Data and code availability statement

The raw data of this study are available upon reasonable request from the authors.

References

- Fokkens WJ, Lund VJ, Hopkins C, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020. *Rhinology* 2020; 58: 1-464.
- 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 Immunol Res* 2020; 12: 176-237.
- 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: 478-484, 484 e471-472.
- Akdis CA, Bachert C, Cingi C, et al. Endotypes and phenotypes of chronic rhinosinusitis: a PRACTALL document of the European Academy of Allergy and Clinical Immunology and the American Academy of Allergy, Asthma & Immunology. *J Allergy Clin Immunol* 2013; 131: 1479-1490.
- Lou H, Zhang N, Bachert C, Zhang L. Highlights of eosinophilic chronic rhinosinusitis with nasal polyps in definition, prognosis, and advancement. *Int Forum Allergy Rhinol* 2018; 8: 1218-1225.
- Liao B, Liu JX, Li ZY, et al. Multidimensional endotypes of chronic rhinosinusitis and their association with treatment outcomes. *Allergy* 2018; 73: 1459-1469.
- Sakuma Y, Ishitoya J, Komatsu M, et al. New clinical diagnostic criteria for eosinophilic chronic rhinosinusitis. *Auris Nasus Larynx* 2011; 38: 583-588.
- Meng Y, Lou H, Wang C, Zhang L. Predictive significance of computed tomography in eosinophilic chronic rhinosinusitis with nasal polyps. *Int Forum Allergy Rhinol* 2016; 6: 812-819.
- Hu Y, Cao PP, Liang GT, Cui YH, Liu Z. Diagnostic significance of blood eosinophil count in eosinophilic chronic rhinosinusitis with nasal polyps in Chinese adults. *Laryngoscope* 2012; 122: 498-503.
- Wu D, Yan B, Wang Y, Zhang L, Wang C. Charcot-Leyden crystal concentration in nasal secretions predicts clinical response to glucocorticoids in patients with chronic rhinosinusitis with nasal polyps. *J Allergy Clin Immunol* 2019; 144: 345-348 e348.
- Tokunaga T, Sakashita M, Haruna T, et al. Novel scoring system and algorithm for classifying chronic rhinosinusitis: the JESREC Study. *Allergy* 2015; 70: 995-1003.
- Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016; 278: 563-577.
- Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017; 14: 749-762.
- Binder ZA, Thorne AH, Bakas S, et al. Epidermal growth factor receptor extracellular domain mutations in glioblastoma present opportunities for clinical imaging and therapeutic development. *Cancer Cell* 2018; 34: 163-177 e167.
- Li X, Liang D, Meng J, et al. Development and validation of a novel computed-tomography enterography radiomic approach for characterization of intestinal fibrosis in crohn's disease. *Gastroenterology* 2021; 160: 2303-2316 e2311.
- Schniering J, Maciukiewicz M, Gabrys HS, et al. Computed tomography-based radiomics decodes prognostic and molecular differences in interstitial lung disease related to systemic sclerosis. *Eur Respir J* 2022; 59.
- Beswick DM, Smith TL, Mace JC, et al. Ethmoid-to-maxillary opacification ratio: a predictor of postoperative olfaction and outcomes in nasal polyposis? *Int Forum Allergy Rhinol* 2021; 11: 48-57.
- van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017; 77: e104-e107.
- Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol* 2016; 34: 2157-2164.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44: 837-845.
- Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol* 2015; 16: e173-180.
- Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; 26: 565-574.
- Kerr KF, McClelland RL, Brown ER, Lumley T. Evaluating the incremental value of new biomarkers with integrated discrimination improvement. *Am J Epidemiol* 2011; 174: 364-374.
- Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008; 27: 157-172; discussion 207-112.
- Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. *Ann Oncol* 2020; 31: 912-920.
- Chow S-C, Shao J, Wang H, Lokhnygina Y. Sample size calculations in clinical research, third edition. 2017.
- Lou H, Meng Y, Piao Y, et al. Cellular phenotyping of chronic rhinosinusitis with nasal

- polyps. *Rhinology* 2016; 54: 150-159.
28. Zhu M, Gao X, Zhu Z, et al. The roles of nasal nitric oxide in diagnosis and endotypes of chronic rhinosinusitis with nasal polyps. *J Otolaryngol Head Neck Surg* 2020; 49: 68.
 29. Lv H, Liu PQ, Xiang R, et al. Predictive and diagnostic value of nasal nitric oxide in eosinophilic chronic rhinosinusitis with nasal polyps. *Int Arch Allergy Immunol* 2020; 181: 853-861.
 30. Sun R, Limkin EJ, Vakalopoulou M, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* 2018; 19: 1180-1191.
 31. Braman N, Prasanna P, Whitney J, et al. Association of peritumoral radiomics with tumor biology and pathologic response to preoperative targeted therapy for HER2 (ERBB2)-positive breast cancer. *JAMA Netw Open* 2019; 2: e192561.
 32. Xu Q, Du K, Zheng M, et al. Application of clinical scores in the differential diagnosis of chronic rhinosinusitis with nasal polyps in a chinese population. *Am J Rhinol Allergy* 2020; 34: 401-408.
 33. Fadda GL, Galizia A, Galizia G, et al. Multiparametric analysis of factors associated with eosinophilic chronic rhinosinusitis with nasal polyps. *Ear Nose Throat J* 2020: 145561320960357.
 34. Yao Y, Zeng M, Liu Z. Revisiting asian chronic rhinosinusitis in the era of type 2 biologics. *Clin Exp Allergy* 2022; 52: 231-243.

Zheng Liu, M.D., Ph.D.

Department of Otolaryngology
 Head and Neck Surgery
 Institute of Allergy and Clinical Immunology
 Tongji Hospital
 Tongji Medical College
 Huazhong University of Science and Technology
 1095 Jiefang Avenue
 Wuhan 430030
 P.R. China

E-mail: zhengliuent@hotmail.com

Or

Ming Zeng, M.D., Ph.D.
 Department of Otolaryngology
 Head and Neck Surgery
 Institute of Allergy and Clinical Immunology
 Tongji Hospital
 Tongji Medical College
 Huazhong University of Science and Technology
 1095 Jiefang Avenue
 Wuhan 430030
 P.R. China

E-mail: zmsx77@163.com

Or

Xiang Lu, M.D., Ph.D.
 Department of Otolaryngology
 Head and Neck Surgery
 Institute of Allergy and Clinical Immunology
 Tongji Hospital
 Tongji Medical College
 Huazhong University of Science and Technology
 1095 Jiefang Avenue
 Wuhan 430030
 P.R. China

mail: luxiangent@hotmail.com

SUPPLEMENTARY MATERIAL

S1. Patients and clinical features

This multicenter study of anonymous data included 3 centers and was approved by Ethics Review Committee at each participating institution. Chronic rhinosinusitis (CRS) with nasal polyps (CRSwNP) was diagnosed according to the international guideline⁽¹⁾. All the patients had ongoing symptoms after initial attempts on medical treatments and underwent endoscopic sinus surgery. The inclusion criteria for patient recruitment were: (i) oral glucocorticoid, and intranasal steroid spray and steroid irrigation treatment were discontinued at least 3 months and 1 month before surgery, respectively; (ii) without acute asthma episode or acute upper airway infection within 1 month before surgery; (iii) not under allergen immunotherapy or biologic treatment. The exclusion criteria for patient recruitment were: (i) fungal sinusitis; (ii) antrochoanal polyps; (iii) cystic fibrosis, vasculitis, primary ciliary dyskinesia, or immunodeficiency; (iv) odontogenic maxillary sinusitis; (v) patients with history of craniocerebral trauma or sinonasal tumours.

Routine blood tests and sinus computed tomography (CT) scans were completed in 1-3 days before surgery. The radiological exclusion criteria were: (i) insufficient CT quality (e.g., metal artifacts, motion artifacts, and lack of intact CT scan of sinuses); (ii) slice thickness more than 2.5 mm.

A total of 431 patients were finally analyzed (Table S1), with 293 from Center 1 (Tongji Hospital), 84 from Center 2 (Xiangyang Central Hospital), and 54 from Center 3 (Shanxi Bethune Hospital). Patients from Center 1 were randomly divided into training cohort (n = 232) and internal validation cohort (n = 61) at a ratio of 8:2, while patients from Center 2 and Center 3 served as external validation cohort-1 and external validation cohort-2, respectively.

The demographic and clinical data including age, sex, comorbidities, previous sinus surgery history, and blood routine test results were retrieved from the medical records anonymously. The symptom score was not included in this study because of its subjectivity. Nasal polyp samples were taken during surgery and subjected to routine hematoxylin and eosin staining. Eosinophilic CRSwNP was diagnosed when the percentage of tissue eosinophils exceeded 10% of the total infiltrating cells as previously described⁽²⁾. Asthma was diagnosed according to the Global Initiative for Asthma (GINA) guidelines⁽³⁾. Allergic rhinitis was diagnosed according to the Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines⁽⁴⁾. Atopic status was evaluated by skin prick test to a standard panel of aeroallergens or by using the ImmunoCAP Phadiatop test (Phadia, Uppsala, Sweden) for detecting immunoglobulin E antibodies against various common inhalant allergens⁽⁵⁾.

There was no significant difference among four study cohorts except that external validation cohort-2 displayed higher fre-

quencies of atopy and allergic rhinitis, higher blood eosinophil/monocyte counts and percentages, and lower anterior ethmoid sinus scores (Table S1). The training cohort, including 93 eosinophilic and 139 noneosinophilic CRSwNP patients, was used to construct the prediction model and determine the optimal parameters. The internal validation cohort including 27 eosinophilic and 34 noneosinophilic CRSwNP patients, external validation cohort-1 including 26 eosinophilic and 58 noneosinophilic CRSwNP patients, and external validation cohort-2 including 28 eosinophilic and 26 noneosinophilic CRSwNP patients, were used to test the prediction model without exposure in the modeling process.

S2. Radiology procedure

Image acquisition

Enrolled patients in the three hospitals underwent standard CT scans using multi-detector row systems. The versions of the instruments and parameters are shown in Table S2. The CT reading was performed on Picture Archiving and Communication Systems (PACS) workstation with adjustable windows, where tri-plane images (axial, coronal and sagittal) were viewed simultaneously and followed by crosshairs for detection and localization of sinus inflammation. The original Digital Imaging and Communications in Medicine (DICOM) images were obtained for 3D reconstruction and subsequent analysis.

Lund–Mackay CT scores

The CT scores were graded by experienced otolaryngologists, who were blinded to the clinical and pathological data. According to Lund–Mackay system⁽¹⁾, the left and right maxillary sinus, anterior and posterior ethmoid sinus, frontal sinus, sphenoid sinus, and osteomeatal complex (OMC) were semi-quantified based on the degree of opacification. Total CT scores were the summary of the scores of left and right sinuses and OMC. The ethmoid-to-maxillary (E/M) ratio was calculated as: (anterior ethmoid sinus score + posterior ethmoid sinus score) / maxillary sinus score⁽⁶⁾.

CT segmentation

The image segmentation was performed by 2 experienced radiologists, who were blinded to clinical and pathological data. As image preprocessing, the images were resampled to isotropic voxels on the 1.25 × 1.25 × 1.25 mm side using a B spline interpolator and normalized to the gray value of the image centered on the image standard deviation value to ensure the consistency of subsequent analyses⁽⁷⁾. Then, region of interest (ROI) was obtained by semi-automatic threshold segmentation method using 3D Slicer (Harvard Medical School, version 4.11.0, <http://www.slicer.org>)⁽⁸⁾, where the main algorithmic functions

used were: (i) brush paint; (ii) threshold segmentation; (iii) level tracing; (iv) median smoothing.

The brush paint was used to outline the entire sinonasal cavity edge as the ROI. The anterior boundary of the nasal cavity was defined as the anterior edge of the nasal bone, while the posterior boundary of the nasal cavity as the choanae, and the other boundaries of the nasal cavity and the borders of all the sinuses were defined as the sinonasal anatomical margins. Threshold segmentation was obtained (visualized by masking) for pixels in the image with values between upper and lower thresholds. The degree of radiation attenuation of high density of bone is different from that of air cavities and soft tissue images. The bone was excluded during threshold segmentation to avoid bias caused by prior surgery in some patients and influence on subsequent feature extraction (e.g., first-order features and texture features). Level tracing interpolated the gray values of pixel points to form a regular grid and got the coordinates and value of each grid point, to estimate the positions of the contour points and further standardize the organization of edge contours. Median smoothing used median interpolation algorithm to remove small details while keeping smooth contours mostly unchanged and set kernel size 2.5 mm to get a more continuous and smooth segmentation result. The CT segmentation was shown in Figure 1A.

Radiomic feature extraction

The feature extraction was implemented with PyRadiomics (version 3.0, <https://pyradiomics.readthedocs.io/en/latest>), an open-source python package for the extraction of radiomic features from medical imaging by referring to the Image Biomarker Standardization Initiative (IBSI) ^(9,10). As showed in Table S3, a total of 1316 radiomic features were extracted from ROI, including: (i) 18 first-order statistics; (ii) 24 shape-based; (iii) 24 Gray Level Cooccurrence Matrix (GLCM); (iv) 16 Gray Level Run Length Matrix (GLRLM); (v) 16 Gray Level Size Zone Matrix (GLSZM); (vi) 5 Neighboring Gray Tone Difference Matrix (NGTDM); (vii) 14 Gray Level Dependence Matrix (GLDM). In addition to the original features, we also extract features from images after wavelet filtration and Laplacian of Gaussian (LoG) filtration. For the wavelet filter, multidimensional convolution can be avoided by applying different weights to the low or high frequency sub-bands. The convolution was carried out with the low (L)-/high (H)-pass “Coiflet 1” wavelet filter along the x-/y-direction, yielding 6 different combinations of decompositions (LHL, LHH, LLH, HLL, HHL and HLH), as compared to the low- or high-frequency sub-bands (LLL and HHH). For the LoG filter, images were filtered using a 3D LoG filter implemented and by changing sigma values to 0.5, 1.0, 1.5, 2.0 and 2.5 mm to suppress noise, yielding another 5 derived images. Here, $\delta = 0.5$ and $\delta = 1.0$ was used to represent fine textures (about 4 pixels or 2.72mm width features), $\delta = 1.5$ and $\delta = 2.0$ were used to represent medium textures (about

6-10 pixels or 4.08-6.8 mm width features), and $\delta = 2.5$ was used to represent coarse textures (about 12 pixels or 8.16 mm width features) ⁽¹¹⁾. Detailed information about the feature names, meanings and mathematical formulas can be obtained from the documentation at <http://pyradiomics.readthedocs.io/en/latest>.

Radiomic feature selection

Features were selected sequentially by using inter-/intra-class correlation coefficients (ICCs) ⁽¹²⁾, and minimum redundancy maximum relevance (mRMR) ⁽¹³⁾. First, the ICCs algorithm, a statistical measure between 0 and 1 indicating null and perfect reproducibility, was used to evaluate inter- and intra-class variation in radiomic feature extraction. Here, we choose the ICC greater than 0.80 as high agreement. To verify inter-class agreement, 30 patients were randomly chosen for manual segmentation by two radiologists (reader-1 and reader-2) to delineate the boundary of each primary ROI, and the radiomic features were extracted. After inter-class ICC calculation and analysis, all segmentations were completed by reader-1. To verify intra-class agreement, three months after the initial segmentation, 50 training datasets were randomly selected and re-segmented by reader-1, then intra-class ICC was calculated. Finally, stable features with both of inter- or intra-class ICC value greater than 0.8 were reserved, and other features were filtered out. Then, mRMR algorithm was used to calculate the mutual information between radiomic features and the endotypes of CRSwNP. The maximum number of predictors should be approximately equal to 1/3 of the smallest population in the training cohort, therefore, the 30 highest ranked features in the mRMR were retained. Radiomic feature selection was shown in Figure 1, A-C.

S3. Model construction

Four types of models were constructed: the CT score model, radiomics model, clinical model, and combined model. In the process of model building within the training cohort, all the validation cohorts were not exposed and only used for the evaluation of the model's performance.

CT score model

The CT score model was constructed using logistic regression (LR) algorithm. Each individual sinus Lund–Mackay CT scores and E/M ratio were analyzed by between-group comparison analysis and univariate analysis to determine the associations with eosinophilic CRSwNP. Then, the CT score features associated with eosinophilic CRSwNP were used to develop the CT score model by LR.

Radiomics model

The radiomics models were constructed using least absolute shrinkage and selection operator (LASSO) and LR algorithm ⁽¹⁴⁾, a widely applicable high-dimensional data regression machine

learning method for radiomics. In the training cohort, we applied the LASSO algorithm to select the best robust predictive features from the 30 extracted radiomic features, and performed penalty parameter tuning by 10-fold cross-validation as showed in Figure 1, D. The LASSO filter retained less than 10 features from the 30 top features generated by mRMR and developed radiomics model by LR. We calculated the respective weighted regression coefficients of the selected features to obtain a rad-score representing the probability of eosinophilic endotype for each patient:

$$\text{Rad - score} = k_1 X_1 + k_2 X_2 + \dots + k_j X_j + b$$

Where X_i is the radiomic feature, k_i is the LR coefficient of X_i and b is the intercept of the LR.

Clinical model

The clinical model was constructed using the LR algorithm. After between-group comparison and univariate analysis, the significant clinical features were analyzed by backward stepwise multivariate LR analysis to identify the independent predictors of eosinophilic CRSwNP. The independent predictors were used to develop the clinical model by LR⁽¹⁵⁾.

Combined model

The combined model was constructed by backward stepwise multivariate LR based on the rad-score and clinical variables. We calculated the respective weighted regression coefficients of the selected variables to obtain a risk-score representing the probability of eosinophilic endotype for each patient:

$$\text{Risk-score} = k_1 \text{clinical - var}_1 + k_2 \text{clinical - var}_2 + \dots + k_j \text{rad-score} + b$$

Where *clinical-var* represents the clinical variable, *rad-score* is calculated in the radiomics model, k_i is the LR coefficient, and b is the intercept of LR.

S4. Model analysis

Discrimination

The discrimination was evaluated through receiver operating characteristic (ROC) with the corresponding area under curve (AUC) value^(16,17). For different models, a larger AUC implies better predictive performance, while a $P < 0.05$ from the DeLong's test implies a significant difference in the predictive effectiveness between two models for the statistical comparison of AUCs. The horizontal axis of ROC represents the false positive rate (FPR = 1- specificity) and the vertical axis of ROC represents the true positive rate (TPR = sensitivity) of the model. An optimal cutoff value was chosen according to the maximum Youden's index in the training cohort.

$$\text{Youden's index} = (\text{sensitivity} + \text{specificity}) - 1$$

Accuracy, sensitivity, specificity, positive predictive value and negative predictive value were calculated.

Calibration

The calibration was evaluated through the Hosmer–Lemeshow test with a calibration curve⁽¹⁸⁾, and $P > 0.05$ means that the model is well calibrated. The horizontal coordinate of the calibration curve represents the probability predicted by our model and the vertical coordinate of the calibration curve represents the actual probability. The closer the slope of the curve is to 1, the better agreement between the predicted outcomes and the real pathological endotypes.

Clinical utility

For prediction models, decision curve analysis (DCA) was used to assess the clinical validity, which integrates and compares the clinical effectiveness of models by calculating the range of net benefits at different threshold probabilities:

$$\text{Net benefit} = (\text{True positives} / n) - (P_t / (1 - P_t)) \times (\text{False positives} / n)$$

Where n is the total number of patients and P_t is the risk threshold. Net benefit was used to assess the benefit that patients can get from the prediction model. The larger net benefit means better clinical utility.

F1 score was used to measure the accuracy of binary classification model. The larger F1 score means the better model, which was calculated as:

$$F_1 = 2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$$

Net reclassification improvement (NRI) value, and integrated discrimination improvement (IDI) value were used to calculate the proportion of patients that were correctly reclassified⁽¹⁹⁾, and to quantify the improvement of diagnostic accuracy of the new model over the original one. NRI and IDI > 0 means the new model has improved in clinical use. The waterfall plots, confusion matrixes and box plots were also performed.

S5. Sample size consideration

To avoid model over-fitting, the number of predictors should be bigger than 1/20–1/8 of the sample size in the training cohort⁽²⁰⁾. In our study, 10 radiomic features were finally selected to build the radiomics model, while the sample size of the training cohort was 232 (with 93 eosinophilic CRSwNP patients). Therefore, the sample size of training cohort was acceptable.

Validation cohorts

The minimum sample size of validation was calculated as fol-

lowing formula ⁽²¹⁾:

Letting the noneosinophilic CRSwNP patients be Group A and eosinophilic CRSwNP patients as Group B, μ represents the mean in each group, and the validation cohorts' minimum sample size and power are calculated respectively:

$$N_A = ((n_A + n_B) / n_B) (\sigma (z_{(1-\alpha/2)} + z_{(1-\beta)}) / (\mu_A - \mu_B))^2$$

$$N_B = ((n_A + n_B) / n_A) (\sigma (z_{(1-\alpha/2)} + z_{(1-\beta)}) / (\mu_A - \mu_B))^2$$

$$1 - \beta = \Phi(z - z_{(1-\alpha/2)}) + \Phi(-z - z_{(1-\alpha/2)}), z = (\mu_A - \mu_B) / (\sigma \sqrt{(1/n_A + 1/n_B)})$$

where, n is the sample size in the training cohort and N is the sample size in the validation cohort, Φ is the standard normal distribution function, α is the Type I error, β is the Type II error, $1 - \beta$ is the power, σ is the variance, while $z_{(1-\alpha/2)}$ and $z_{(1-\beta)}$ are obtained by consulting the normal distribution table.

Under the desired two-sided significance level of $\alpha = 0.05$ and power of $1 - \beta = 0.95$, the sample size needed in the validation cohort was calculated to be 23 for noneosinophilic CRSwNP and 16 for eosinophilic CRSwNP. All the validation cohorts met the sample size requirements.

S6. Statistical analysis

All statistical analyses were performed using Python software (The Python Software Foundation, version 3.7.4) and R software (R Foundation for Statistical Computing, version 4.2.1). All the packages used in this study are detailed in Table E4. Data distri-

bution was tested for normality using the Kolmogorov-Smirnov test. For continuous variables, t-/t'-test or Mann-Whitney U test was used to analyze the difference between 2 groups depending on data normality, while ANOVA test or Kruskal-Wallis H test was used to assess significant among more than 2 groups. For categorical variables, Fisher's exact test or chi-square test/chi-square test with Yates' was used to compare the difference between groups.

S7. Radiomics score (rad-score) calculation formula

$$\begin{aligned} \text{Rad-score} = & (-0.173 \times \text{Wavelet (LLH)}_Firstorder_Mean) + \\ & (-0.095 \times \text{Wavelet (HLL)}_GLCM_IMC2) + \\ & (-0.043 \times \text{Original_Firstorder_Minimum}) + \\ & (-0.034 \times \text{Wavelet (LHL)}_Firstorder_Skewness) + \\ & (-0.020 \times \text{Original_Shape_Flatness}) + \\ & (0.054 \times \text{Wavelet (HHH)}_Firstorder_Median) + \\ & (0.127 \times \text{Wavelet (HLH)}_Firstorder_Mean) + \\ & (0.192 \times \text{Original_Firstorder_Maximum}) + \\ & (0.429 \times \text{Wavelet (HHH)}_Firstorder_Mean) + \\ & (0.460 \times \text{Wavelet (HLH)}_GLCM_IMC1) - 0.451 \end{aligned}$$

Among them, we divided these 10 features into 3 categories and provide potential interpretable clinically relevant information: A) first-order features; B) shape feature; C) texture features.

The code used for analysis were uploaded on GitHub (https://github.com/Yukeyuzhang/CRS_Radiomics.git). The raw data of this study are available upon reasonable request from the authors.

References

- Fokkens WJ, Lund VJ, Hopkins C, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020. *Rhinology* 2020; 58: 1-464.
- 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: 478-484, 484 e471-472.
- Brozek JL, Bousquet J, Agache I, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines-2016 revision. *J Allergy Clin Immunol* 2017; 140: 950-958.
- Reddel HK, Bacharier LB, Bateman ED, et al. Global Initiative for Asthma Strategy 2021: Executive Summary and Rationale for Key Changes. *Respirology* 2022; 27: 14-35.
- Pradaliere A. The diagnosis of allergic diseases. *Rev Prat* 2007; 57: 1313-1319.
- Beswick DM, Smith TL, Mace JC, et al. Ethmoid-to-maxillary opacification ratio: a predictor of postoperative olfaction and outcomes in nasal polyposis? *Int Forum Allergy Rhinol* 2021; 11: 48-57.
- Shafiq-Ul-Hassan M, Zhang GG, Latifi K, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* 2017; 44: 1050-1062.
- van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017; 77: e104-e107.
- Zwanenburg A, Vallières M, Abdallah MA, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology* 2020; 295: 328-338.
- Bettinelli A, Marturano F, Avanzo M, et al. A novel benchmarking approach to assess the agreement among radiomic tools. *Radiology* 2022; 303: E30.
- Ng F, Ganeshan B, Kozarski R, Miles KA, Goh V. Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. *Radiology* 2013; 266: 177-184.
- Barry WT, Kernagis DN, Dressman HK, et al. Intratumor heterogeneity and precision of microarray-based predictors of breast cancer biology and clinical outcome. *J Clin Oncol* 2010; 28: 2198-2206.
- De Jay N, Papillon-Cavanagh S, Olsen C, et al. mRMRe: an R package for parallelized mRMR ensemble feature selection. *Bioinformatics* 2013; 29: 2365-2368.
- Yang J, Song X, Lai Y, et al. Development and validation of a postoperative nomogram for predicting overall survival after endoscopic surgical management of olfactory neuroblastoma. *EClinicalMedicine* 2020; 29-30: 100577.
- Zheng J, Yu H, Batur J, et al. A multicenter study to develop a non-invasive radiomic model to identify urinary infection stone in vivo using machine-learning. *Kidney Int* 2021; 100: 870-880.
- Massion PP, Antic S, Ather S, et al. Assessing the accuracy of a deep learning method to risk stratify indeterminate pulmonary nodules. *Am J Respir Crit Care Med* 2020; 202: 241-249.
- Kammer MN, Lakhani DA, Balar AB, et al. Integrated biomarkers for the management of indeterminate pulmonary nodules. *Am J Respir Crit Care Med* 2021; 204: 1306-1316.
- Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA* 2017; 318: 1377-1384.
- Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; 26: 565-574.

20. Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. *Ann Oncol* 2020; 31: 912-920. third edition. 2017.
21. Chow S-C, Shao J, Wang H, Lokhnygina Y. *Sample size calculations in clinical research*,

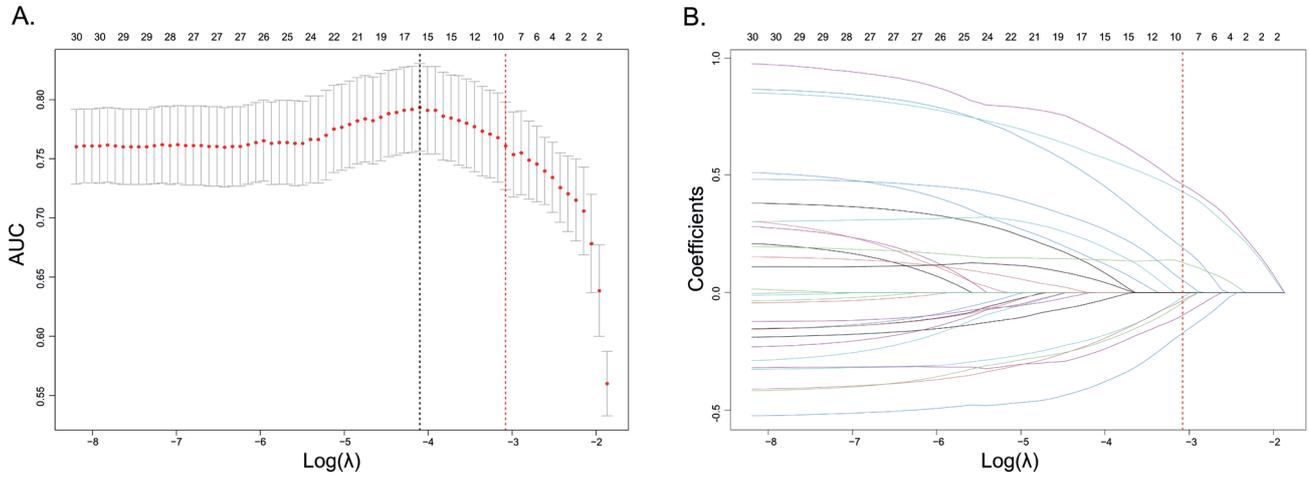


Figure S1. Basolateral viral release over time of Calu-3 cells (A) and pNECs (B).

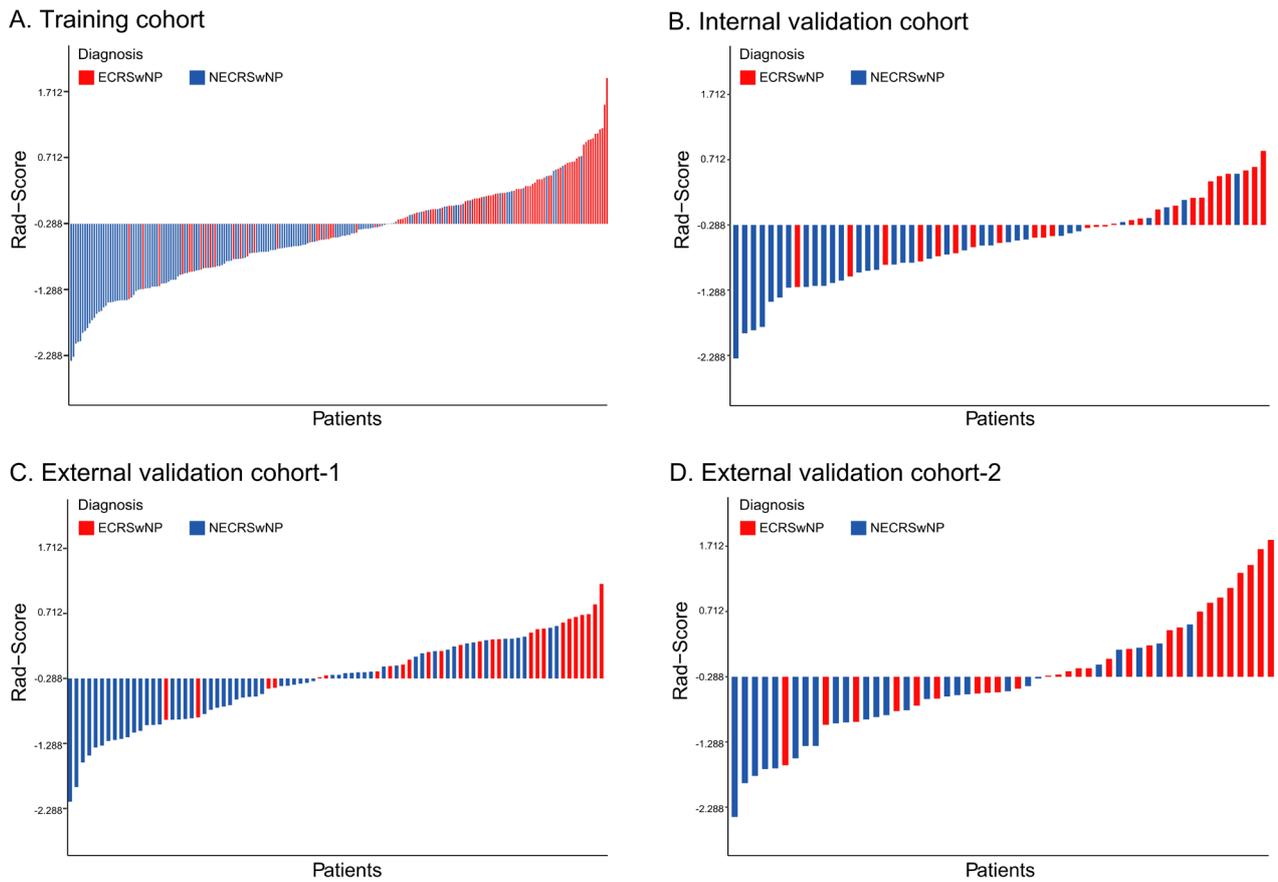


Figure S2. Rad-score for each patient in every cohort. The horizontal coordinate represents each patient and the vertical coordinate represents the rad-score calculated from the radiomic characteristics of each patient through the LASSO-LR method. ECRSwNP patients are represented in red, while NECRSwNP patients in blue.

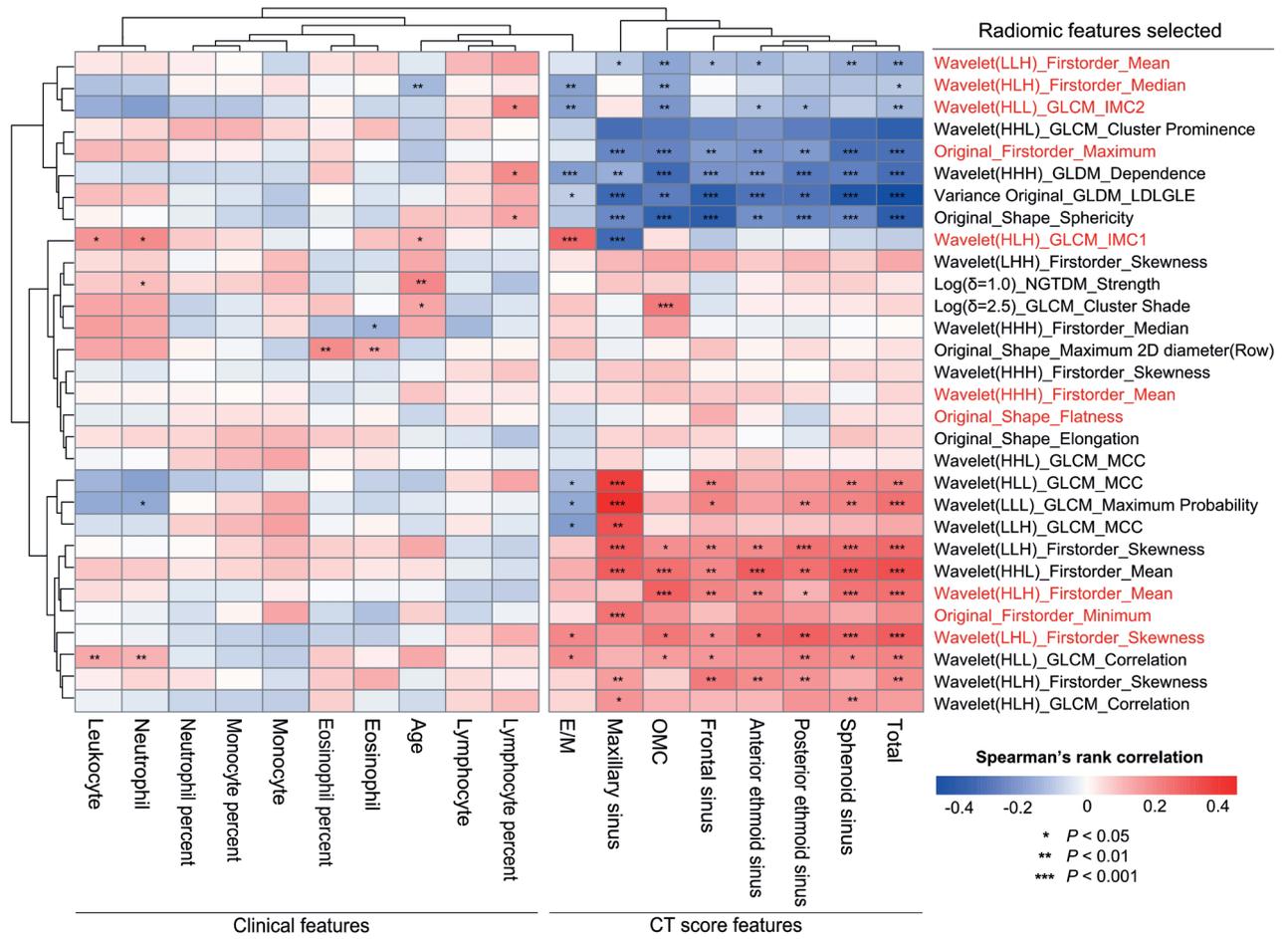
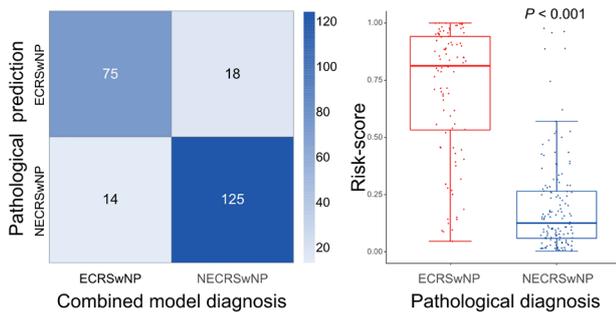
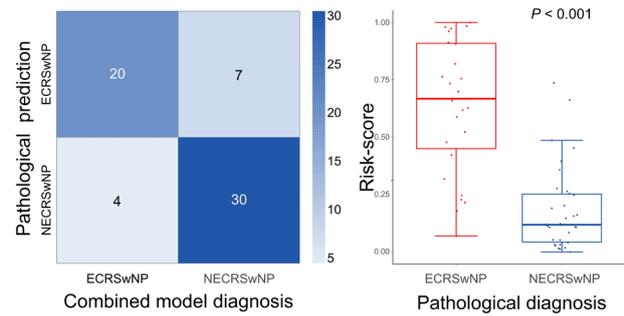


Figure S3. The heatmap of Spearman's correlation between radiomic features, Lund-Mackay CT scores and clinical features in the training cohort. The color scheme and P value symbols showing the degree of relevance are displayed in the lower right corner. The 30 radiomics features on the right were selected by mRMR, and the 10 features in red colour were selected by LASSO from 30 features.

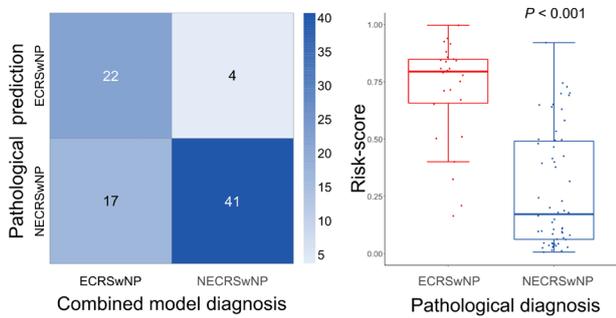
A. Training cohort



B. Internal validation cohort



C. External validation cohort-1



D. External validation cohort-2

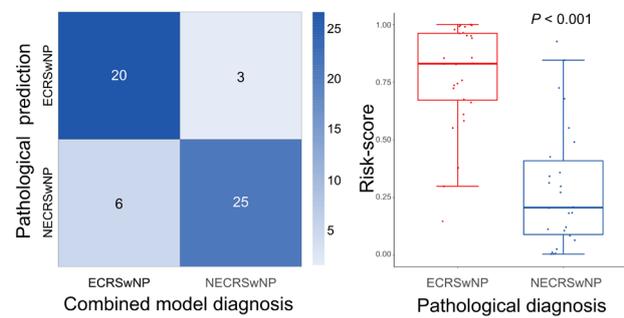


Figure S4. Confusion matrix and box plots demonstrating the evaluation of the combined model for discriminating eosinophilic and noneosinophilic CRSwNP in every cohort. For the confusion matrix, the horizontal axis represents the actual pathological diagnosis of the patient, while the vertical axis represents the prediction by the combined model, and the colour scheme on the right indicates the numbers of patients. For the box plot, the horizontal axis represents the actual pathological diagnosis of the patient, and the vertical axis represents the risk-scores calculated by the combined model.

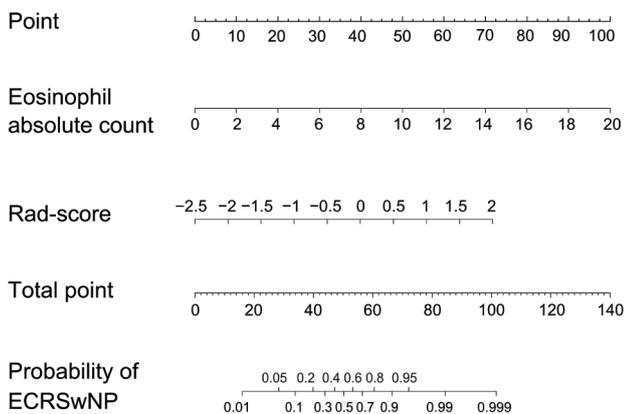


Figure S5. The nomogram developed on the basis of the combined model.

Table S1. Demographic and clinical characteristics of patients in different study cohorts.

Characteristics	Training cohort (n = 232)	Internal validation cohort (n = 61)	External validation cohort-1 (n = 84)	External validation cohort-2 (n = 54)	P*
ECRSwNP, n (%)	93 (40.09)	27 (44.26)	26 (30.95)	28 (51.85)	0.093
Age, median (IQR), years	45.0 (28.0-54.0)	42.0 (27.5-54.5)	50.0 (41.0-56.0)	43.5 (28.8-55.3)	0.060
Sex, male, n (%)	154 (66.38)	43 (70.49)	61 (72.62)	40 (74.07)	0.578
Environment, city, n (%)	186 (80.17)	48 (78.69)	60 (71.43)	36 (66.67)	0.111
Smoking habits, n (%)	46 (19.83)	12 (19.67)	18 (21.43)	13 (24.07)	0.908
Drinking habits, n (%)	41 (17.67)	10 (16.39)	21 (25.00)	13 (24.07)	0.367
Prior sinus surgery, n (%)	57 (24.57)	12 (19.67)	12 (14.29)	13 (24.07)	0.246
Allergic rhinitis, n (%)	51 (21.98)	12 (19.67)	15 (17.86)	22 (40.74)	0.010
Asthma, n (%)	36 (15.52)	11 (18.03)	11 (13.10)	11 (20.37)	0.677
Atopy, n (%)	66 (28.45)	14 (22.95)	15 (17.86)	25 (46.30)	0.003
Blood routine examination					
Leukocyte absolute count, median (IQR), $\times 10^9/L$	6.25 (5.38-7.45)	6.41 (5.31-7.70)	5.76 (4.98-6.89)	6.55 (4.95-7.93)	0.148
Neutrophil absolute count, median (IQR), $\times 10^9/L$	3.30 (2.70-4.15)	3.26 (2.70-4.14)	3.15 (2.54-4.10)	3.45 (2.53-4.31)	0.728
Neutrophil percent, median (IQR), %	52.60 (47.73-59.30)	54.40 (46.80-60.50)	55.60 (48.80-60.63)	53.45 (48.95-57.73)	0.629
Lymphocyte absolute count, median (IQR), $\times 10^9/L$	2.06 (1.68-2.48)	2.23 (1.68-2.50)	2.04 (1.57-2.53)	1.97 (1.55-2.51)	0.697
Lymphocyte percent, median (IQR), %	33.45 (28.43-39.55)	32.00 (28.70-38.60)	34.70 (28.58-39.23)	32.80 (28.45-37.33)	0.557
Eosinophil absolute count, median (IQR), $\times 10^9/L$	0.18 (0.10-0.37)	0.19 (0.10-0.46)	0.16 (0.09-0.31)	0.40 (0.18-0.56)	0.002
Eosinophil percent, median (IQR), %	3.00 (1.60-6.00)	2.90 (1.55-6.70)	2.65 (1.63-4.98)	6.30 (2.65-8.83)	0.003
Monocyte absolute count, median (IQR), $\times 10^9/L$	0.45 (0.36-0.56)	0.47 (0.34-0.63)	0.35 (0.30-0.44)	0.45 (0.30-0.57)	<0.001
Monocyte percent, median (IQR), %	7.35 (6.26-8.90)	7.20 (6.00-8.60)	5.80 (5.20-7.38)	6.85 (5.30-7.83)	<0.001
CT Score, median (IQR)					
Frontal sinus	2 (0-3)	1 (0-3.5)	2 (0.25-4)	2 (0-3)	0.420
Anterior ethmoid sinus	2 (2-4)	2 (2-3)	2 (2-3)	2 (2-3)	0.002
Posterior ethmoid sinus	2 (2-3)	2 (1-2)	2 (2-2)	2 (1.75-3)	0.395
Maxillary sinus	2 (2-3)	2 (2-2)	2 (2-3)	2 (2-2.25)	0.149
Sphenoid sinus	1 (0-2)	1 (0-2)	1 (0-2)	1 (0-2)	0.924
OMC	4 (2-4)	4 (2-4)	4 (2-4)	4 (2-4)	0.326
Total CT score	13 (8-17)	13 (7.5-17)	13.5 (10-18)	12 (8.75-17.25)	0.659
E/M	2 (1.7-2.5)	2 (1.5-2.17)	2 (1.3-2)	2 (1.46-2.5)	0.108

Abbreviations: ECRSwNP, eosinophilic chronic rhinosinusitis with nasal polyps; IQR, interquartile range; OMC, ostiomeatal complex; E/M, (anterior ethmoid sinus score + posterior ethmoid sinus score) / (maxillary sinus score).

Data are presented as median and interquartile ranges and were analyzed by Kruskal-Wallis H test. *The P values in red color indicate those less than 0.05.

Table S2. The image acquisition parameters in every medical center.

Parameters		Center 1: Wuhan Tongji Hospital		Center 2: Xiangyang Central Hospital	Center 3: Shanxi Bethune Hospital
CT version	Spectral CT (Discovery CT 750 HD scanner, GE Healthcare, USA)	Spectral CT (Brilliance iCT, ROYAL PHILIPS, Netherlands)	Spectral CT (SOMATOM Definition AS+, Siemens Healthineers, Germany)	Spectral CT (uCT510 scanner, United Imaging, China)	Spectral CT (SOMATON sensation64, Siemens Healthineers, Germany)
CT tube voltage	Spectral imaging mode switching between 120 kVp and 140 kVp	120kVp	120kVp	120kVp	120kVp
Image matrix	512×512	512×512	512×512	512×512	512×512
Field of view	500×500 mm	500×500 mm	500×500 mm	500×500 mm	500×500 mm
Reconstruction image thickness	0.625 mm	1.25 mm	2.5 mm	1.0 mm	0.625 mm

Table S3. Extracted radiomic features.

Group	Subgroup	Radiomic features
Firstorder features	--	Energy, Total Energy, Entropy, Minimum, 10th percentile, 90th percentile, Maximum, Mean, Median, Interquartile Range, Range, Mean Absolute Deviation (MAD), Robust Mean Absolute Deviation (rMAD), Root Mean Squared (RMS), Skewness, Kurtosis, Variance, Uniformity
Shape features	--	Mesh Volume, Voxel Volume, Surface Area, Surface Area to Volume ratio, Sphericity, Maximum 2D diameter (Slice), Maximum 2D diameter (Column), Maximum 2D diameter (Row), Major Axis Length, Minor Axis Length, Least Axis Length, Elongation, Flatness, Maximum 3D diameter
Texture features	Gray Level Cooccurrence Matrix (GLCM)	Autocorrelation, Joint Average, Cluster Prominence, Cluster Shade, Cluster Tendency, Contrast, Correlation, Difference Average, Difference Entropy, Difference Variance, Joint Energy, Joint Entropy, Informational Measure of Correlation (IMC) 1, Informational Measure of Correlation (IMC) 2, Inverse Difference Moment (IDM), Maximal Correlation Coefficient (MCC), Inverse Difference Moment Normalized (IDMN), Inverse Difference (ID), Inverse Difference Normalized (IDN), Inverse Variance, Maximum Probability, Sum Average, Sum Entropy, Sum of Squares
	Gray Level Run Length Matrix (GLRLM)	Short Run Emphasis (SRE), Long Run Emphasis (LRE), Gray Level Non-Uniformity (GLN), Gray Level Non-Uniformity Normalized (GLNN), Run Length Non-Uniformity (RLN), Run Length Non-Uniformity Normalized (RLNN), Run Percentage (RP), Gray Level Variance (GLV), Run Variance (RV), Run Entropy (RE), Low Gray Level Run Emphasis (LGLRE), High Gray Level Run Emphasis (HGLRE), Short Run Low Gray Level Emphasis (SRLGLE), Short Run High Gray Level Emphasis (SRHGLE), Long Run Low Gray Level Emphasis (LRLGLE), Long Run High Gray Level Emphasis (LRHGLE)
	Gray Level Size Zone Matrix (GLSZM)	Small Area Emphasis (SAE), Large Area Emphasis (LAE), Gray Level Non-Uniformity (GLN), Gray Level Non-Uniformity Normalized (GLNN), Size- Zone Non-Uniformity (SZN), Size-Zone Non-Uniformity Normalized (SZNN), Zone Percentage (ZP), Gray Level Variance (GLV), Zone Variance (ZV), Zone Entropy (ZE), Low Gray Level Zone Emphasis (LGLZE), High Gray Level Zone Emphasis (HGLZE), Small Area Low Gray Level Emphasis (SALGLE), Small Area High Gray Level Emphasis (SAHGLE), Large Area Low Gray Level Emphasis (LALGLE), Large Area High Gray Level Emphasis (LAHGLE)
	Neighboring Gray Tone Difference Matrix (NGTDM)	Coarseness, Contrast, Busyness, Complexity, Strength
	Gray Level Dependence Matrix (GLDM)	Small Dependence Emphasis (SDE), Large Dependence Emphasis (LDE), Gray Level Non-Uniformity (GLN), Dependence Non-Uniformity (DN), Dependence Non-Uniformity Normalized (DNN), Gray Level Variance (GLV), Dependence Variance (DV), Dependence Entropy (DE), Low Gray Level Emphasis (LGLE), High Gray Level Emphasis (HGLE), Small Dependence Low Gray Level Emphasis (SDLGLE), Small Dependence High Gray Level Emphasis (SDHGLE), Large Dependence Low Gray Level Emphasis (LDLGLE), Large Dependence High Gray Level Emphasis (LDHGLE)
LoG filtered Features *	--	LoG ($\delta = 0.5$) _x , LoG ($\delta = 1.0$) _x , LoG ($\delta = 1.5$) _x , LoG ($\delta = 2.0$) _x , LoG ($\delta = 2.5$) _x
Wavelet features *	--	Wavelet (LLL) _x , Wavelet (LLH) _x , Wavelet (LHL) _x , Wavelet (LHH) _x , Wavelet (HLL) _x , Wavelet (HLH) _x , Wavelet (HHL) _x , Wavelet (HHH) _x

* x denotes the first-order statistics features and the statistics-based textural features listed above.

Table S4. Packages and functions used.

Software	Purposes	Packages	Functions	Versions
Python	Data import	"Path"	Path ()	15.1.2
	Data analysis, time series, and statistics	"pandas"	pandas.read_csv (), pandas.read_excel (), pandas.concat ()	1.2.4
	Array computing	"numpy"	np.zeros (), np.hstack ()	1.20.1
	Statistical ICC	"pingouin"	pg.intraclass_corr ()	0.4.0
R	Data import	"readxl"	read.table (), read.excel (), write.table ()	1.3.1
	Statistical calculations and random number generation	"stats"	chisq.test (), shapiro.test (), var.test (), t.test (), wilcox.test (), cor (), coef (), predict (), glm (), quantile (), sd (), hclust ()	4.1.1
	Parallelized mRMR ensemble feature selection	"mRMR"	mRMR.data (), mRMR.ensemble ()	2.1.2
	Random samples and permutations	"base"	Sample (), table (), list (), mean ()	4.1.1
	Fit a GLM with lasso or elastic net regularization	"glmnet"	cv.glmnet (), glmnet ()	4.1.2
	Visualizing the Performance of Scoring Classifiers	"ROCR", "pROC"	Prediction (), performance (), plot.roc ()	1.0.11, 1.18.0
	Streamline the model training process	"caret"	Sensitivity (), specificity (), ConfusionMatrix ()	6.0.9
	Composer of Plots	"ggplot2", "cowplot", "patchwork"	Ggplot (), patchwork ()	3.3.5, 1.1.1, 1.1.1
	Transformation functions and fitting functions	"rms", "MASS"	Datadist (), lrm (), nomogram (), calibrate (), stepAIC ()	6.2.0
	Risk Model Decision Analysis	"rmda"	decision_curve (), plot_decision_curve ()	1.6
	Reclassification table and statistics	"PredictABEL"	Reclassification (), PlotCalibration ()	1.2.4
	NRI and IDI to models	"nricens"	Nricens ()	1.6

Table S5. The rad-score in eosinophilic and noneosinophilic CRSwNP stratified by prior surgery history in all enrolled patients.

	Rad-score, median (IQR)		P*
	Eosinophilic	Non-eosinophilic	
No prior surgery (n = 305)	0.129 (-0.506 to 0.523)	-0.652 (-1.206 to -0.236)	<0.001
With prior surgery (n = 87)	0.002 (-0.466 to 0.446)	-0.715 (-1.195 to -0.331)	<0.001

*The P values in red color indicate those less than 0.05.

Table S6. Logistic regression analysis of clinical features associated with eosinophilic CRSwNP in the training cohort

	Univariate logistic regression		Multivariate logistic regression	
	OR (95% CI)	P*	OR (95% CI)	P*
Eosinophil absolute count	1.917 (1.592 to 2.362)	<0.001	1.917 (1.592 to 2.362)	<0.001
Eosinophil percent	1.433 (1.286 to 1.616)	<0.001	--	--
Asthma	2.767 (1.345 to 5.860)	0.006	--	--

*The P values in red color indicate those less than 0.05.

Table S7. Reclassification analysis of combined model in four cohorts

	Combined model over clinical model		Combined model over radiomics model	
	NRI (95% CI, P*)	IDI (95% CI, P*)	NRI (95% CI, P*)	IDI (95% CI, P*)
Training cohort	0.862 (0.656-1.068, P < 0.001)	0.229 (0.173-0.284, P < 0.001)	0.805 (0.560-1.009, P < 0.001)	0.209 (0.153-0.264, P < 0.001)
Internal validation cohort	1.092 (0.708-1.475, P < 0.001)	0.321 (0.200-0.442, P < 0.001)	1.092 (0.712-1.471, P < 0.001)	0.289 (0.172-0.405, P < 0.001)
External validation cohort-1	0.639 (0.261-1.017, P < 0.001)	0.175 (0.077-0.272, P < 0.001)	0.678 (0.295-1.061, P < 0.001)	0.179 (0.079-0.279, P < 0.001)
External validation cohort-2	1.003 (0.594-1.411, P < 0.001)	0.275 (0.156-0.394, P < 0.001)	0.967 (0.559-1.375, P < 0.001)	0.248 (0.131-0.366, P < 0.001)

*The P values in red color indicate those less than 0.05.