

Smoking status stratifies Th 1 /Th 2 phenotypes in pre-COPD high-risk individuals: quantile-specific blood biomarkers for primary care

Xiaohua Zhang

Shanghai Fifth People's Hospital, Fudan University

Juan Xie

Shanghai Fifth People's Hospital, Fudan University

Shuang Yang

Jiangchuan Community Health Service Center of Minhang District

Biying Wu

Shanghai Fifth People's Hospital, Fudan University

Jindong Shi

Pujiang Hospital of Shanghai

Xiaoying Hu

`huxiaoying633@126.com`

Jiangchuan Community Health Service Center of Minhang District

Research Article

Keywords: Pre-COPD, High-risk population, Smoking status, Th1/Th2 polarization, Quantile regression, Primary care

Posted Date: July 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10083621/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Smoking status stratifies Th₁/Th₂ phenotypes in pre-COPD high-risk individuals: quantile-specific blood biomarkers for primary care

Authors

Xiaohua Zhang^{1,5,†}, Juan Xie^{1,5,†}, Shuang Yang^{2,5}, Biying Wu⁶, Jindong Shi^{3,4,5*},
Xiaoying Hu^{2,5*}

[†]These authors contributed equally to this work.

*Corresponding author.

Email: Xiaoying Hu, huxiaoying633@126.com

Jindong Shi, shijindong@fudan.edu.cn

Affiliations

¹Department of General Medicine, Shanghai Fifth People's Hospital, Fudan University, Shanghai, China

²Department of General Medicine, Jiangchuan Community Health Service Center of Minhang District, Shanghai, China

³Department of Respiratory Medicine, Pujiang Hospital of Shanghai, Shanghai, China

⁴Department of Pulmonary and Critical Care Medicine, Shanghai Fifth People's Hospital, Fudan University, Shanghai, China

⁵Center for Community-Based Health Research, Fudan University, Shanghai, China

⁶Department of Laboratory Medicine, Shanghai Fifth People's Hospital, Fudan University, Shanghai, China

Abstract

Background Chronic obstructive pulmonary disease (COPD) is most modifiable at the pre-COPD high-risk stage, yet early immune dysregulation remains poorly characterized, limiting the development of primary care-based early screening tools. While Th₁/Th₂ phenotypes are well-described in established COPD, their quantile-specific distributions in community pre-COPD individuals, particularly when stratified by smoking status, remain not systematically characterized.

Methods This community-based cross-sectional study enrolled 149 participants aged ≥ 40 years. Participants were classified into healthy, high-risk and COPD groups according to the 2024 Chinese primary care COPD guidelines (combined COPD-SQ questionnaire and post-bronchodilator spirometry), and further stratified by smoking status. Quantile regression at $q = 0.25, 0.50$ and 0.75 was performed to identify group differences after adjusting for age, sex, and smoking index (smoking subgroup only).

Results Smoking status stratified distinct Th₁/Th₂ phenotypes primarily in the pre-COPD high-risk group. Relative to healthy controls, smoking-related pre-COPD individuals had Th₁-dominant immunity with isolated low-quantile TNF- α elevation ($\beta = 1.100$, 95% CI: 0.321–1.879), accompanied by neutrophilia, lymphopenia, reduced MCHC and elevated RDW-CV. Never-smoking pre-COPD individuals showed Th₂-skewed polarization, with mid-quantile IL-5 elevation ($\beta = 1.376$, 95% CI: 0.200–2.553), decreased IL-12p70, and universal MCHC reduction. These early stage-specific biomarkers normalized in established COPD, where smoking patients had elevated IL-6/IL-8 and never-smoking patients exhibited increased eosinophils.

Conclusion Smoking status stratifies distinct Th₁/Th₂ phenotypes in pre-COPD high-risk individuals. Quantile-specific peripheral blood biomarkers can detect early immune dysregulation missed by conventional mean regression, and may provide preliminary, cost-effective indicators for stratified pre-COPD screening in primary care.

Keywords: Pre-COPD; High-risk population; Smoking status; Th₁/Th₂ polarization; Quantile regression; Primary care

1 Introduction

Chronic obstructive pulmonary disease (COPD) is a major global public health issue associated with high morbidity and mortality^[1-3]. Its early clinical manifestations are insidious, and most patients are diagnosed only after irreversible airflow limitation has developed, resulting in delayed intervention and an increased socioeconomic burden^[3]. Pre-COPD, a high-risk stage characterized by early symptoms and structural changes, therefore represents a critical window for effective prevention and early intervention.

As the primary risk factor for COPD, smoking drives distinct inflammatory endotypes in established disease: smokers typically exhibit Th₁-dominant neutrophilic inflammation, whereas non-smokers often present with Th₂-skewed eosinophilic phenotypes^[4-6]. Although systemic inflammation is recognized as a core pathological mechanism underlying COPD progression^[4, 7], critical gaps remain in our understanding of inflammatory heterogeneity during the pre-COPD high-risk stage. While Th₁/Th₂ polarization is well-characterized in diagnosed patients, it remains unclear whether these phenotypes emerge before airflow limitation occurs, and how smoking status modulates early immune dysregulation.

Spirometry is the gold standard for COPD diagnosis, but its widespread application in primary care is severely limited by equipment shortages and lack of trained personnel^[8]. The COPD Screening Questionnaire (COPD-SQ), a validated screening tool originally developed for Chinese populations^[9] and widely applied in primary care screening settings^[10], enables rapid community screening but suffers from a high false-positive rate when used alone. The 2024 Chinese Primary Care COPD Guidelines therefore recommend a combined strategy of COPD-SQ plus post-bronchodilator spirometry to accurately identify high-risk individuals^[11]. However, no studies to date have systematically characterized smoking-stratified biological profiles in this dual-criteria defined pre-COPD population.

Routine blood parameters and inflammatory cytokines are cost-effective, non-invasive, and widely accessible biomarkers that reflect systemic inflammation and oxidative stress, making them ideal for primary care-based screening^[6, 7, 12-14].

Nevertheless, most studies have focused exclusively on established COPD rather than pre-COPD, seldom considered smoking as a critical effect modifier of inflammatory phenotypes^[6, 15], and relied almost entirely on conventional mean regression, which inherently fails to capture subtle, quantile-specific early abnormalities restricted to specific segments of the biomarker distribution^[5].

Accordingly, we enrolled community-dwelling adults and classified them into healthy, high-risk and COPD groups using the guideline-recommended combined criteria^[11]. After stratification by smoking status and confounder adjustment, quantile regression was used to explore quantile-specific changes in routine blood and inflammatory markers. This exploratory study aimed to delineate smoking-related Th₁/Th₂ inflammatory phenotypes in pre-COPD high-risk individuals and identify quantile-specific peripheral blood biomarkers that can support stratified early screening in primary care settings.

2 Materials and Methods

2.1 Study Design and Participants

A community-based cross-sectional study was conducted using stratified cluster sampling among residents aged ≥ 40 years in Jiangchuan Subdistrict, Minhang District, Shanghai, from August 2023 to July 2024. The study protocol was approved by the Ethics Committee of Shanghai Fifth People's Hospital, Fudan University (Approval No. 2023-153). This observational cross-sectional study was not registered as a clinical trial; clinical trial number: not applicable. Written informed consent was obtained from all participants prior to any study-related procedures.

Inclusion criteria:

- (1) Age ≥ 40 years;
- (2) Local community residence for ≥ 1 year;
- (3) Ability to complete standardized questionnaires, spirometry testing, and venous blood sampling.

Exclusion criteria:

- (1) Severe acute illness or surgical intervention within the preceding 3 months;

- (2) Uncontrolled hypertension (systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg despite medication) or diabetes mellitus (fasting blood glucose ≥ 7.0 mmol/L or random blood glucose ≥ 11.1 mmol/L);
- (3) Acute respiratory tract infection within the preceding 4 weeks;
- (4) Congenital or acquired immunodeficiency, connective tissue disease, hematological disease, or active malignant tumor;
- (5) Use of glucocorticoids, immunosuppressive agents, or non-steroidal anti-inflammatory drugs (NSAIDs) within the preceding 1 month;
- (6) Contraindications to spirometry or inability to cooperate with spirometry testing.

2.2 Baseline Assessment and Grouping

2.2.1 Baseline Assessment

Baseline data were collected by well-trained primary care physicians using a standardized questionnaire. Collected variables included age, sex, educational level, occupation, medical history, medication history, COPD-SQ score^[9, 10], smoking status, and smoking index.

Smoking status was classified into three categories:

- (1) Never-smokers: Individuals who had smoked < 100 cigarettes in their lifetime;
- (2) Current smokers: Individuals who had smoked ≥ 1 cigarette per day for ≥ 6 months and were still smoking at the time of enrollment;
- (3) Former smokers: Individuals who had quit smoking for ≥ 6 months before enrollment.

For the purpose of this study, current and former smokers were combined into the smoking subgroup for analysis, as both groups have an increased risk of COPD. Smoking index was calculated as packs per day $\times$ years of smoking (1 pack = 20 cigarettes).

Spirometry was performed using a calibrated PF680 spirometer (Yiliankang Medical Technology Co., Ltd., Zhejiang, China) in accordance with the *Guidelines for Pulmonary Function Tests (2022 Edition)*. Pre-bronchodilator forced expiratory volume in 1 second (FEV₁), forced vital capacity (FVC), and FEV₁/FVC ratio were measured. For participants with pre-bronchodilator FEV₁/FVC < 0.7 ,

post-bronchodilator FEV₁/FVC was assessed 15 minutes after inhalation of 400 µg salbutamol. All tests were performed by certified technicians, with at least 3 acceptable curves (variability between the two highest FEV₁ values < 5% or < 0.15 L) and the optimal value used for analysis. Daily calibration of the spirometer was performed using a 3-L syringe to ensure measurement accuracy.

2.2.2 Grouping

Participants were classified into three groups according to the *2024 Chinese Guidelines for Primary Care Diagnosis and Management of COPD*^[11], based on COPD-SQ score and post-bronchodilator spirometry:

- (1) Healthy group: COPD-SQ score < 16 and post-bronchodilator FEV₁/FVC ≥ 0.7;
- (2) Pre-COPD high-risk group: COPD-SQ score ≥ 16 and post-bronchodilator FEV₁/FVC ≥ 0.7;
- (3) COPD group: Post-bronchodilator FEV₁/FVC < 0.7 (regardless of COPD-SQ score).

All participants were further stratified into smoking subgroup (current + former smokers) and never-smoking subgroup for subgroup analysis.

2.3 Laboratory Measurements

Fasting peripheral venous blood (5 mL) was collected from all participants between 7:00 and 9:00 AM after an overnight fast of ≥8 hours. Samples were immediately aliquoted into two EDTA anticoagulant tubes. One tube was used directly for complete blood count analysis within 2 hours of collection. The second tube was centrifuged at 3000 × g for 10 minutes at 4°C within 30 minutes of venipuncture to separate plasma, which was then divided into 200 µL aliquots and stored at -80°C until analysis. All samples were thawed only once to avoid cytokine degradation.

2.3.1 Complete Blood Count Measurements

Complete blood count parameters were measured using a fully automated hematology analyzer (BC-5180CRP; Mindray Medical International Ltd., Shenzhen, China). Measured parameters included white blood cell count, neutrophil count and percentage, lymphocyte count and percentage, eosinophil (EOS) count and percentage, basophil count and percentage, monocyte count and percentage, red blood cell count,

hemoglobin (Hb), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red blood cell distribution width–standard deviation (RDW-SD), red blood cell distribution width–coefficient of variation (RDW-CV), platelet count, platelet distribution width (PDW), mean platelet volume (MPV), large platelet ratio, and plateletcrit (PCT). All assays were conducted strictly following the manufacturer’s protocols. Daily quality control and instrument calibration were conducted to ensure analytical precision and reproducibility.

2.3.2 Inflammatory Cytokine Measurement

Plasma levels of 12 inflammatory cytokines were quantified using a multiplex bead-based flow cytometry immunoassay with a RaiseCyte 2L6C flow cytometer (RaiseCare Bio-Tech Co., Ltd., Qingdao, China). A commercial 12-plex cytokine assay kit (Cat. No. BNCBA002-96T; Saihan Bio-Tech Co., Ltd., Shanghai, China) was used to detect interleukin-5 (IL-5), interferon- α (IFN- α), IL-2, IL-6, IL-1 β , IL-10, IFN- γ , IL-8, IL-17, IL-4, IL-12p70, and tumor necrosis factor- α (TNF- α). All experiments were conducted strictly in accordance with the manufacturer’s protocols. Blank controls, standard curves and internal quality control samples were included in each assay run to guarantee analytical accuracy, precision and reproducibility.

2.4 Statistical Analysis

Data were analyzed using IBM SPSS Statistics 27.0, and all graphical visualizations were generated using GraphPad Prism 10.0. A two-sided $P < 0.05$ was considered statistically significant.

The Shapiro–Wilk test was used to evaluate the normality of continuous variables. Since most variables exhibited non-normal distributions, continuous data were presented as median (interquartile range) [M(P25, P75)], and categorical variables were expressed as counts and percentages [n(%)].

A two-way nonparametric interaction model based on rank transformation was applied to assess the interactive effect between smoking status and study grouping.

Categorical variables were analyzed using the Pearson χ^2 test or Fisher’s exact test, as appropriate. Intergroup comparisons of continuous variables were performed using

the Kruskal-Wallis H test, with post-hoc pairwise comparisons conducted via the Mann-Whitney U test followed by Bonferroni correction for multiple testing. Only biomarkers with significant between-group differences ($P < 0.05$) in each subgroup were included in subsequent quantile regression analyses to reduce multiple comparison burden.

Quantile regression was conducted at the 25th, 50th, and 75th percentiles ($q = 0.25, 0.50, 0.75$) to explore the independent effect of study group after adjusting for age, sex, and smoking index (only in the smoking subgroup). Model goodness-of-fit was evaluated using pseudo R^2 and mean absolute error (MAE).

Spearman's rank correlation analysis was performed to assess the associations between smoking index and core hematological and inflammatory biomarkers in the total study population and the smoking subgroup.

3 Results

3.1 Baseline characteristics of study participants

To ensure adequate statistical power for between-group comparisons, we used a targeted recruitment strategy based on community electronic health records. Residents with a history of abnormal spirometry or documented COPD risk factors (smoking, chronic cough, etc.) were preferentially invited, along with age- and sex-matched healthy residents.

A total of 180 eligible residents were initially enrolled. Thirty-one participants were excluded according to predefined criteria, leaving 149 participants in the final analysis. These were classified into healthy controls ($n=45$), pre-COPD high-risk individuals ($n=56$), and established COPD patients ($n=48$) (Fig. 1). All participants were further stratified into smoking ($n=67$) and never-smoking ($n=82$) subgroups.

Baseline characteristics are presented in Table 1. In the overall study population, no significant between-group differences were observed in age, sex, smoking status, or smoking index (all $P > 0.05$). Similarly, age and sex were comparable among the three groups in both the smoking and never-smoking subgroups, and smoking index was balanced across groups in the smoking subgroup (all $P > 0.05$).

Table 1. Baseline characteristics of study participants overall and by smoking status

Characteristic	Healthy (n=45)	Pre-COPD (n=56)	COPD (n=48)	Statistic	P-value
Overall population					
Age, years	66.0 (60.5, 69.0)	66.0 (63.0, 70.0)	67.0 (63.0, 70.0)	H=1.090	0.580
Sex, n (%)				$\chi^2=0.312$	0.855
Male	24 (53.3)	30 (53.6)	28 (58.3)		
Female	21 (46.7)	26 (46.4)	20 (41.7)		
Smoking status, n (%)				$\chi^2=2.908$	0.234
Never-smoker	25 (55.6)	35 (62.5)	22 (45.8)		
Smoker	20 (44.4)	21 (37.5)	26 (54.2)		
Smoking index, pack-years	0.0 (0.0, 28.5)	0.0 (0.0, 24.1)	8.5 (0.0, 42.0)	H=4.224	0.121
Smoking subgroup	Healthy (n=20)	Pre-COPD (n=21)	COPD (n=26)		
Age, years	67.5 (61.0, 74.0)	68.0 (64.5, 72.0)	67.0 (63.5, 69.3)	H=0.361	0.835
Sex, n (%)				—	0.197
Male	20 (100.0)	21 (100.0)	24 (92.3)		
Female	0 (0.0)	0 (0.0)	2 (7.7)		
Smoking index, pack-years	29.0 (23.8, 39.8)	26.5 (21.8, 38.5)	41.0 (19.5, 48.3)	H=1.875	0.392
Never-smoking subgroup	Healthy (n=25)	Pre-COPD (n=35)	COPD (n=22)		
Age, years	64.0 (60.5, 68.5)	65.0 (63.0, 70.0)	66.5 (63.0, 71.0)	H=3.280	0.194
Sex, n (%)				—	0.620
Male	4 (16.0)	9 (25.7)	4 (18.2)		
Female	21 (84.0)	26 (74.3)	18 (81.8)		
Smoking index, pack-years	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	H=0.000	1.000

Footnotes:

Continuous variables are presented as median (interquartile range, P25, P75) and compared using the Kruskal–Wallis H test.

Categorical variables are presented as number (percentage) and compared using the Pearson χ^2 test (overall) or Fisher's exact test (smoking subgroups). A *P*-value < 0.05 was considered statistically significant. Smoking index = packs of cigarettes per day × years of smoking (1 pack = 20 cigarettes). Smoking index was zero in all never-smoking participants.

3.2 Biomarker screening and preliminary between-group comparisons

3.2.1 Interaction between smoking status and study group

Two-way nonparametric tests based on rank transformation were performed to assess the interactive effect between smoking status and study grouping. Significant smoking status-by-group interactions were detected for neutrophil count and percentage, lymphocyte percentage, EOS count, Hb, HCT, MCV, RDW-CV, IL-2, IL-1 β , and IFN- γ (all *P* < 0.05). No significant interactions were observed for white blood cell count, lymphocyte count, EOS percentage, MCHC, IL-5, IL-6, IL-8, IL-4, TNF- α or any other measured indicators (all *P* > 0.05). These results indicate that smoking status acts as a key effect modifier for hematological and inflammatory biomarkers, justifying subsequent stratified analyses by smoking status.

3.2.2 Biomarker screening and preliminary comparisons in the smoking subgroup

In the smoking subgroup, Kruskal-Wallis H tests identified significant overall

between-group differences in 10 biomarkers (all $P < 0.05$), which were selected for subsequent quantile regression analysis: 7 routine blood parameters (neutrophil count and percentage, lymphocyte percentage, Hb, MCH, MCHC, RDW-CV) and 3 inflammatory cytokines (IL-6, IL-8, TNF- α).

Post-hoc pairwise comparisons showed that pre-COPD high-risk individuals had higher neutrophil counts and percentages, lower lymphocyte percentages, Hb, MCH and MCHC, and elevated RDW-CV compared with healthy controls. Notably, RDW-CV in the high-risk group was significantly higher than in both healthy controls and COPD patients. COPD patients had significantly higher IL-6 and IL-8 levels than healthy controls (Table 2). No other significant pairwise differences were detected.

Table 2. Biomarker screening and preliminary comparisons in the smoking subgroup

Parameter	Healthy (n=20)	Pre-COPD (n=21)	COPD (n=26)	H	P value	Pairwise comparison: Z statistic/adjusted P value
Neutrophil percentage (%)	59.00 (53.85, 64.53)	64.60 (60.25, 71.60)	67.75 (61.33, 71.60)	12.288	0.002	Pre-COPD vs H: 2.692/ 0.021 Pre-COPD vs C: -0.520/1.000 C vs H: 3.341/ 0.003
Lymphocyte percentage (%)	32.30 (26.38, 36.68)	25.30 (21.60, 30.95)	23.70 (19.98, 28.90)	13.460	0.001	Pre-COPD vs H: -2.652/ 0.024 Pre-COPD vs C: 0.788/1.000 C vs H: -3.564/ 0.001
Neutrophil count ($\times 10^9/L$)	3.62 (3.00, 4.31)	4.40 (4.26, 5.13)	4.81 (3.85, 5.64)	14.764	<0.001	Pre-COPD vs H: 3.168/ 0.005 Pre-COPD vs C: -0.212/1.000 C vs H: 3.537/ 0.001
Hb (g/L)	158.00 (155.00, 166.75)	144.00 (137.00, 158.00)	153.50 (135.75, 161.00)	7.815	0.020	Pre-COPD vs H: -2.673/ 0.023 Pre-COPD vs C: -0.690/1.000 C vs H: -2.128/0.100
MCH (pg)	31.65 (30.30, 32.10)	30.60 (28.80, 31.30)	31.00 (29.68, 31.80)	8.251	0.016	Pre-COPD vs H: -2.865/ 0.013 Pre-COPD vs C: -1.670/0.285 C vs H: -1.362/0.519
MCHC (g/L)	336.50 (333.25, 343.00)	329.00 (320.00, 332.50)	335.50 (325.00, 339.25)	8.213	0.016	Pre-COPD vs H: -2.863/ 0.013 Pre-COPD vs C: -1.597/0.331 C vs H: -1.432/0.457
RDW-CV (%)	12.45 (11.75, 12.70)	13.00 (12.70, 13.50)	12.35 (12.08, 13.03)	14.325	<0.001	Pre-COPD vs H: 3.663/ 0.001 Pre-COPD vs C: 2.717/ 0.020 C vs H: -1.167/0.730
IL-6 (pg/mL)	2.24 (1.37, 3.57)	3.11 (1.98, 5.56)	4.62 (2.37, 6.20)	8.633	0.013	Pre-COPD vs H: 1.791/0.220 Pre-COPD vs C: -1.060/0.868 C vs H: 2.927/ 0.010
IL-8 (pg/mL)	3.89 (1.83, 7.99)	4.12 (2.08, 6.21)	6.90 (3.96, 14.26)	7.653	0.022	Pre-COPD vs H: 0.266/1.000 Pre-COPD vs C: -2.214/0.080 C vs H: 2.464/ 0.041
TNF- α (pg/mL)	1.46 (1.03, 3.15)	2.81 (2.30, 5.64)	2.20 (1.57, 6.30)	6.684	0.035	Pre-COPD vs H: 2.583/ 0.029 Pre-COPD vs C: 1.254/0.630 C vs H: 1.477/0.419

Footnotes:

1. Continuous variables are presented as median (interquartile range) [M (P25, P75)].
 2. Overall intergroup comparisons were performed using the Kruskal–Wallis H test. A P-value < 0.05 was considered statistically significant.
 3. Pairwise comparisons were performed using the Mann–Whitney U test with Bonferroni correction for parameters with significant overall differences.
- Results are expressed as Z statistic / adjusted P value.

4. Abbreviations: H, healthy group; Pre-COPD, pre-COPD high-risk group; C, COPD group; Hb, hemoglobin; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW-CV, red blood cell distribution width–coefficient of variation; IL, interleukin; TNF- α , tumor necrosis factor- α ;
5. Only parameters with significant intergroup differences are presented.

3.2.3 Biomarker screening and preliminary comparisons in the never-smoking subgroup

In the never-smoking subgroup, Kruskal-Wallis H tests identified significant overall between-group differences in 10 biomarkers (all $P<0.05$), which were selected for subsequent quantile regression analysis: 6 routine blood parameters (EOS count and percentage, HCT, MCV, MCHC, PDW) and 4 inflammatory cytokines (IL-5, IL-2, IFN- γ , IL-12p70).

Post-hoc tests showed that pre-COPD high-risk individuals had higher MCV and lower MCHC than healthy controls, and MCHC was also significantly reduced compared with the COPD group. COPD patients exhibited significantly higher EOS counts and percentages than both the high-risk and healthy groups, as well as lower PDW than healthy controls. For inflammatory cytokines, the high-risk group had significantly higher IL-5 and lower IL-12p70 than healthy controls. COPD patients had markedly lower IL-2, IFN- γ and IL-12p70 than healthy controls, while IFN- γ levels were higher in the high-risk group than in the COPD group (Table 3). No other significant pairwise differences were observed.

Table 3. Biomarker screening and preliminary comparisons in the never-smoking subgroup

Parameter	Healthy (n=25)	Pre-COPD (n=35)	COPD (n=22)	H	P value	Pairwise comparison: Z statistic/adjusted P value
EOS percentage (%)	2.50 (1.35, 3.55)	2.30 (1.40, 3.60)	3.95 (2.70, 4.83)	8.441	0.015	Pre-COPD vs H: 0.171/1.000 Pre-COPD vs C: -2.588/ 0.029 C vs H: 2.562/ 0.031
EOS count ($\times 10^9/L$)	0.16 (0.09, 0.24)	0.13 (0.08, 0.21)	0.24 (0.15, 0.29)	8.028	0.018	Pre-COPD vs H: 0.209/1.000 Pre-COPD vs C: -2.505/ 0.037 C vs H: 2.518/ 0.035
HCT (%)	40.80 (38.90, 43.65)	44.00 (41.20, 46.50)	41.25 (38.55, 44.58)	6.329	0.042	Pre-COPD vs H: 2.378/0.052 Pre-COPD vs C: 1.706/0.264 C vs H: 0.542/1.000
MCV (fL)	91.10 (87.70, 94.70)	94.30 (90.80, 96.40)	92.00 (88.80, 94.65)	6.647	0.036	Pre-COPD vs H: 2.462/ 0.041 Pre-COPD vs C: 1.687/0.275 C vs H: 0.635/1.000
MCHC (g/L)	335.00 (328.50, 341.50)	323.00 (320.00, 329.00)	331.50 (322.75, 337.25)	16.563	<0.001	Pre-COPD vs H: -3.954/ 0.000 Pre-COPD vs C: -2.468/ 0.041 C vs H: -1.245/0.640
PDW (fL)	16.10 (16.00, 16.30)	16.10 (15.80, 16.20)	15.80 (15.23, 16.20)	6.644	0.036	Pre-COPD vs H: -1.076/0.846

Parameter	Healthy (n=25)	Pre-COPD (n=35)	COPD (n=22)	H	P value	Pairwise comparison: Z statistic/adjusted P value
						Pre-COPD vs C: 1.714/0.260 C vs H: -2.559/ 0.031
IL-5 (pg/mL)	1.72 (1.48, 2.95)	3.09 (1.83, 4.98)	1.57 (1.22, 4.08)	7.466	0.024	Pre-COPD vs H: 2.437/ 0.044 Pre-COPD vs C: 2.109/0.105 C vs H: -0.221/1.000
IL-2 (pg/mL)	1.77 (1.62, 2.12)	1.61 (1.27, 2.14)	1.42 (0.91, 1.81)	6.190	0.045	Pre-COPD vs H: -1.543/0.369 Pre-COPD vs C: 1.169/0.727 C vs H: -2.470/ 0.040
IFN- γ (pg/mL)	5.00 (3.47, 6.34)	4.21 (3.75, 6.01)	2.84 (1.73, 4.59)	8.048	0.018	Pre-COPD vs H: -0.305/1.000 Pre-COPD vs C: 2.462/ 0.041 C vs H: -2.564/ 0.031
IL-12p70 (pg/mL)	1.90 (1.78, 2.12)	1.53 (1.23, 1.80)	1.32 (1.00, 1.70)	14.910	<0.001	Pre-COPD vs H: -2.985/ 0.008 Pre-COPD vs C: 1.046/0.886 C vs H: -3.648/ 0.001

Footnotes:

- Continuous variables are presented as median (interquartile range) [M (P25, P75)].
- Overall intergroup comparisons were performed using the Kruskal–Wallis H test. A P value < 0.05 was considered statistically significant.
- Pairwise comparisons were performed using the Mann–Whitney U test with Bonferroni correction for parameters with significant overall differences. Results are expressed as Z statistic / adjusted P value.
- Abbreviations: H, healthy group; Pre-COPD, pre-COPD high-risk group; C, COPD group; EOS, eosinophil; HCT, hematocrit; MCV, mean corpuscular volume; MCHC, mean corpuscular hemoglobin concentration; PDW, platelet distribution width; IL, interleukin; IFN- γ , interferon- γ .
- Only parameters with significant intergroup differences are presented.

3.3 Quantile regression analysis of subgroup-specific significant biomarkers

3.3.1 Quantile regression in the smoking subgroup

After adjustment for age, sex and smoking index, quantile regression was conducted at $q = 0.25, 0.50$ and 0.75 . All models showed adequate goodness-of-fit (Table 4).

For routine blood indicators, the pre-COPD high-risk group had higher neutrophil percentage at $q = 0.25$ and $q = 0.50$ than the healthy group ($\beta = 8.076$, 95% CI: 1.436–14.715; $\beta = 5.620$, 95% CI: 0.412–10.829), and neutrophil count was also significantly elevated across all three quantiles (all $P < 0.05$). Lymphocyte percentage was significantly reduced in the high-risk group at $q = 0.25$ and $q = 0.50$ ($\beta = -4.222$, 95% CI: -8.235–-0.209; $\beta = -6.972$, 95% CI: -12.818–-1.126). The high-risk group also exhibited lower Hb ($q = 0.25, 0.50$), lower MCH ($q = 0.25$), lower MCHC ($q = 0.25, 0.50$) and higher RDW-CV ($q = 0.25, 0.50$) compared with the healthy group (all $P < 0.05$), and showed decreased MCH/MCHC and increased RDW-CV relative to the COPD group at multiple quantiles (all $P < 0.05$).

For inflammatory cytokines, IL-6 at $q = 0.50$ and IL-8 at $q = 0.25$ were significantly higher in the COPD group than in the healthy group ($\beta = 1.885$, 95% CI: 0.489–3.281;

$\beta = 2.859$, 95% CI: 0.451–5.267). Notably, TNF- α was significantly elevated only at $q=0.25$ in the high-risk group compared with both healthy controls ($\beta = 1.100$, 95% CI: 0.321–1.879) and the COPD group ($\beta = 0.899$, 95% CI: 0.137–1.661), with no significant differences at higher quantiles.

Quantile-specific variations in biomarkers for the smoking subgroup are presented in Fig. 2.

Table 4. Quantile regression in the smoking subgroup (adjusted for age, sex, and smoking index)

Outcome Variable	Quantile	Model Fit Indices (Pseudo R ² /MAE)	Pre-COPD vs Healthy β (95%CI)	COPD vs Healthy β (95%CI)	Pre-COPD vs COPD β (95%CI)
Neutrophil percentage (%)	0.25	0.129/6.9723	8.076 (1.436, 14.715)*	7.384 (0.837, 13.932)*	0.691 (−5.806, 7.188)
	0.50	0.150/5.6189	5.620 (0.412, 10.829)*	6.518 (1.382, 11.654)*	−0.898 (−5.994, 4.198)
	0.75	0.211/6.4489	3.615 (−1.728, 8.957)	7.561 (2.293, 12.830)*	−3.946 (−9.174, 1.281)
Lymphocyte percentage (%)	0.25	0.215/6.4992	−4.222 (−8.235, −0.209)*	−7.870 (−11.827, −3.912)*	3.647 (−0.280, 7.574)
	0.50	0.132/5.5672	−6.972 (−12.818, −1.126)*	−6.293 (−12.058, −0.528)*	−0.679 (−7.064, 5.706)
	0.75	0.132/6.8197	−4.889 (−11.691, 1.913)	−6.522 (−13.230, 0.186)	1.633 (−5.023, 8.289)
Neutrophil count (×10 ⁹ /L)	0.25	0.271/0.7597	0.908 (0.516, 1.299)*	0.719 (0.333, 1.105)*	0.409 (0.014, 0.804)*
	0.50	0.191/0.6145	0.956 (0.325, 1.587)*	1.366 (0.744, 1.989)*	−0.410 (−1.028, 0.208)
	0.75	0.195/0.8103	0.813 (0.018, 1.608)*	1.394 (0.610, 2.178)*	−0.581 (−1.320, 0.158)
Hb (g/L)	0.25	0.210/11.7648	−17.174 (−33.698, −0.649)*	−7.000 (−23.296, 9.296)	−7.020 (−21.873, 7.833)
	0.50	0.155/10.7703	−13.375 (−23.180, −3.570)*	−8.357 (−18.026, 1.312)	−6.709 (−15.985, 2.568)
	0.75	0.104/12.7482	−10.337 (−23.689, 3.015)	−7.852 (−21.019, 5.315)	−6.467 (−19.829, 6.894)
MCH (pg)	0.25	0.102/1.4427	−2.093 (−3.410, −0.776)*	−0.763 (−2.062, 0.536)	−1.512 (−2.704, −0.319)*
	0.50	0.102/1.1525	−0.739 (−1.931, 0.454)	−0.461 (−1.637, 0.715)	−0.278 (−1.445, 0.889)
	0.75	0.106/1.3410	−0.739 (−1.931, 0.454)	−0.461 (−1.637, 0.715)	−0.557 (−1.742, 0.627)
MCHC (g/L)	0.25	0.119/9.6123	−13.557 (−20.569, −6.545)*	−7.589 (−14.504, −0.674)*	−5.968 (−12.829, 0.893)
	0.50	0.091/7.6668	−8.933 (−16.515, −1.352)*	−0.689 (−8.165, 6.788)	−8.244 (−15.773, −0.716)*
	0.75	0.056/8.5301	−9.640 (−20.099, 0.819)	−1.523 (−11.837, 8.791)	−8.117 (−18.351, 2.118)
RDW-CV (%)	0.25	0.081/1.4457	1.117 (0.627, 1.606)*	0.302 (−0.181, 0.785)	0.815 (0.336, 1.294)*
	0.50	0.058/1.3589	0.488 (0.020, 0.955)*	−0.142 (−0.603, 0.319)	0.630 (0.172, 1.087)*
	0.75	0.042/1.4742	0.580 (−0.248, 1.408)	0.130 (−0.686, 0.947)	0.450 (−0.360, 1.260)
IL-6 (pg/mL)	0.25	0.106/1.6711	0.719 (−0.515, 1.952)	1.163 (−0.053, 2.380)	−0.444 (−1.651, 0.762)
	0.50	0.239/1.3736	0.648 (−0.767, 2.064)	1.885 (0.489, 3.281)*	−1.262 (−2.634, 0.109)
	0.75	0.178/1.7744	1.076 (−1.614, 3.766)	2.081 (−0.571, 4.734)	−1.006 (−3.638, 1.626)
IL-8 (pg/mL)	0.25	0.088/4.4728	0.501 (−1.941, 2.942)	2.859 (0.451, 5.267)*	−2.359 (−4.748, 0.031)
	0.50	0.125/3.9354	0.894 (−3.062, 4.850)	2.483 (−1.418, 6.384)	−1.589 (−5.460, 2.282)
	0.75	0.166/4.8169	−0.128 (−9.555, 9.299)	4.642 (−4.654, 13.938)	−4.770 (−14.206, 4.666)
TNF- α (pg/mL)	0.25	0.092/2.6892	1.100 (0.321, 1.879)*	0.201 (−0.567, 0.969)	0.899 (0.137, 1.661)*
	0.50	0.095/2.4500	0.918 (−1.280, 3.116)	1.172 (−0.996, 3.340)	−0.253 (−2.404, 1.897)
	0.75	0.216/2.7940	2.021 (−2.439, 6.480)	0.681 (−3.718, 5.079)	−4.770 (−14.206, 4.666)

Footnotes:

1. Data were presented as regression coefficients (β) and 95% confidence intervals (95%CI); models adjusted for age, sex, and smoking index.
2. Model fit indices: Pseudo R², model explanatory power (higher = better fit); MAE, mean absolute error (lower = higher prediction accuracy).
3. *Statistically significant (95% CI does not include 0).

4. Abbreviations: Hb: hemoglobin; MCH: mean corpuscular hemoglobin; MCHC: mean corpuscular hemoglobin concentration; RDW-CV: red cell distribution width-coefficient of variation; IL: interleukin; TNF- α : tumor necrosis factor- α .

3.3.2 Quantile regression in the never-smoking subgroup

After adjustment for age and sex, all quantile regression models showed adequate goodness-of-fit (Table 5).

For routine blood indicators, no significant differences in EOS percentage and count were detected between the pre-COPD high-risk group and healthy group across all quantiles (all $P > 0.05$). In contrast, EOS levels were substantially higher in the COPD group than in the healthy and high-risk groups at most quantiles (all $P < 0.05$). Notably, the high-risk group exhibited significantly reduced MCHC relative to the healthy group at all three quantiles ($q = 0.25$: $\beta = -10.000$, 95% CI: -15.643 – -4.357 ; $q = 0.50$: $\beta = -11.500$, 95% CI: -17.091 – -5.909 ; $q = 0.75$: $\beta = -12.267$, 95% CI: -21.915 – -2.619), and MCHC was also lower than in the COPD group at $q = 0.50$.

For inflammatory cytokines, IL-5 was significantly increased in the high-risk group at $q = 0.50$ relative to the healthy group ($\beta = 1.376$, 95% CI: 0.200 – 2.553), and was also higher than in the COPD group at $q = 0.25$ and $q = 0.50$ ($\beta = 0.615$, 95% CI: 0.195 – 1.035 ; $\beta = 1.494$, 95% CI: 0.281 – 2.707). In addition, IL-12p70 was markedly lower in both the high-risk and COPD groups than in the healthy group at $q = 0.25$ and $q = 0.50$ (all $P < 0.05$), and remained reduced in the COPD group at $q = 0.75$ ($\beta = -0.463$, 95% CI: -0.891 – -0.035).

Quantile-specific variations in biomarkers for the never-smoking subgroup are displayed in Fig. 3.

Table 5. Quantile regression in the never-smoking subgroup (adjusted for age and sex)

Outcome Variable	Quantile	Model Fit Indices (Pseudo R ² /MAE)	High-risk vs Healthy β (95%CI)	COPD vs Healthy β (95%CI)	High-risk vs COPD β (95%CI)
EOS percentage (%)	0.25	0.072/1.5183	0.200 (−0.530, 0.930)	1.400 (0.578, 2.222)*	−1.200 (−1.953, −0.447)*
	0.50	0.099/1.2814	0.300 (−0.699, 1.299)	2.175 (1.051, 3.299)*	−2.125 (−3.180, −1.070)*
	0.75	0.110/1.6096	−0.137 (−1.465, 1.190)	1.469 (−0.026, 2.963)	−1.606 (−2.975, −0.237)*
EOS count (×10 ⁹ /L)	0.25	0.055/0.1033	−0.015 (−0.067, 0.038)	0.068 (0.009, 0.127)*	−0.083 (−0.137, −0.028)*
	0.50	0.072/0.0860	0.029 (−0.042, 0.099)	0.106 (0.027, 0.185)*	−0.078 (−0.150, −0.005)*
	0.75	0.091/0.1052	0.020 (−0.074, 0.115)	0.124 (0.018, 0.230)*	−0.104 (−0.201, −0.006)*

Outcome Variable	Quantile	Model Fit Indices (Pseudo R ² /MAE)	High-risk vs Healthy β (95%CI)	COPD vs Healthy β (95%CI)	High-risk vs COPD β (95%CI)
HCT (%)	0.25	0.086/4.2584	3.444 (−0.044, 6.933)	0.044 (−3.883, 3.972)	3.400 (−0.198, 6.998)
	0.50	0.062/3.5473	1.840 (−0.814, 4.494)	0.360 (−2.628, 3.348)	3.000 (0.121, 5.879)*
	0.75	0.083/4.1646	2.967 (−1.990, 7.923)	2.133 (−3.447, 7.713)	0.833 (−4.278, 5.944)
MCV (fL)	0.25	0.061/3.8063	2.900 (−0.366, 6.166)	1.060 (−2.617, 4.737)	1.840 (−1.528, 5.208)
	0.50	0.083/3.0232	2.400 (−0.535, 5.335)	0.650 (−2.654, 3.954)	1.750 (−1.277, 4.777)
	0.75	0.074/3.6451	1.450 (−0.857, 3.757)	−0.475 (−3.072, 2.122)	2.400 (−0.256, 5.056)
MCHC (g/L)	0.25	0.144/7.9756	−10.000 (−15.643, −4.357)*	−5.750 (−12.102, 0.602)	−4.250 (−10.068, 1.568)
	0.50	0.163/6.8882	−11.500 (−17.091, −5.909)*	−3.667 (−9.960, 2.627)	−7.833 (−13.598, −2.069)*
	0.75	0.154/8.5236	−12.267 (−21.915, −2.619)*	−3.467 (−14.327, 7.394)	−9.000 (−19.268, 1.268)
PDW (fL)	0.25	0.055/0.7038	−0.125 (−1.337, 1.087)	−0.662 (−2.026, 0.701)	0.538 (−0.712, 1.787)
	0.50	0.033/0.6309	−0.075 (−0.359, 0.209)	−0.383 (−0.703, −0.064)*	0.308 (0.016, 0.601)*
	0.75	0.010/0.6805	−0.100 (−0.366, 0.166)	−0.100 (−0.399, 0.199)	0.000 (−0.274, 0.274)
IL-5 (pg/mL)	0.25	0.045/1.7006	0.305 (−0.102, 0.712)	−0.310 (−0.768, 0.149)	0.615 (0.195, 1.035)*
	0.50	0.076/1.5407	1.376 (0.200, 2.553)*	−0.118 (−1.442, 1.207)	1.494 (0.281, 2.707)*
	0.75	0.060/2.0933	1.391 (−1.191, 3.973)	0.552 (−2.354, 3.459)	0.839 (−1.824, 3.973)
IL-2 (pg/mL)	0.25	0.062/0.8283	−0.370 (−0.751, 0.011)	−0.693 (−1.122, −0.264)*	0.323 (−0.070, 0.716)
	0.50	0.029/0.7339	−0.156 (−0.488, 0.177)	−0.236 (−0.610, 0.139)	0.090 (−0.252, 0.433)
	0.75	0.025/0.8750	−0.043 (−1.105, 1.019)	−0.520 (−1.716, 0.675)	0.477 (−0.618, 1.573)
IFN- γ (pg/mL)	0.25	0.141/1.8604	0.435 (−0.645, 1.515)	−1.424 (−2.639, −0.208)*	1.858 (0.744, 2.972)*
	0.50	0.125/1.5226	−0.496 (−1.683, 0.692)	−2.329 (−3.665, −0.992)*	1.833 (0.609, 3.057)*
	0.75	0.100/1.9515	−0.317 (−2.383, 1.750)	−2.177 (−4.503, 0.150)	1.860 (−0.271, 3.991)
IL-12p70 (pg/mL)	0.25	0.142/0.4444	−0.570 (−0.889, −0.252)*	−0.764 (−1.122, −0.405)*	0.193 (−0.135, 0.522)
	0.50	0.149/0.3906	−0.435 (−0.718, −0.151)*	−0.552 (−0.871, −0.233)*	0.117 (−0.175, 0.410)
	0.75	0.065/0.4758	−0.273 (−0.653, 0.107)	−0.463 (−0.891, −0.035)*	0.190 (−0.202, 0.582)

Footnotes:

1. Data were presented as regression coefficients (β) and 95% confidence intervals (95%CI); models adjusted for age and gender.
2. Model fit indices: Pseudo R², model explanatory power (higher = better fit); MAE, mean absolute error (lower = higher prediction accuracy).
3. *Statistically significant (95% CI does not include 0).
4. Abbreviations: EOS, eosinophil; HCT, hematocrit; MCV, mean corpuscular volume; MCHC, mean corpuscular hemoglobin concentration; PDW, platelet distribution width; IL, interleukin; IFN- γ , interferon- γ .

3.4 Correlation between smoking index and core biomarkers

To explore the dose-response relationship between cumulative smoking exposure and inflammatory phenotypes, we performed Spearman’s rank correlation analysis between smoking index and the core biomarkers identified in the above analyses. Correlation patterns are presented in the heatmap (Fig. 4), where red indicates positive correlations, blue indicates negative correlations, and asterisks indicate statistical significance ($P < 0.05$).

In the total study population, smoking index was positively correlated with neutrophil percentage ($r_s = 0.361$, $P < 0.001$), neutrophil count ($r_s = 0.326$, $P < 0.001$), Hb ($r_s = 0.353$, $P < 0.001$) and MCHC ($r_s = 0.194$, $P = 0.018$), and negatively correlated with lymphocyte percentage ($r_s = -0.343$, $P < 0.001$). For inflammatory cytokines,

smoking index was positively associated with IL-6, IL-8 and TNF- α , and negatively associated with IL-12p70 (all $P < 0.05$).

In the smoking subgroup, all significant associations between smoking index and routine blood indicators were abolished (all $P > 0.05$). However, the positive correlations between smoking index and Th₁-related cytokines were strengthened: IL-6 ($r_s = 0.375$, $P = 0.002$) and TNF- α ($r_s = 0.323$, $P = 0.008$). IL-8 showed a marginally significant correlation ($r_s = 0.235$, $P = 0.056$), and no other significant associations were detected (all $P > 0.05$).

Discussion

This exploratory community-based cross-sectional study is, to our knowledge, among the first studies to apply quantile regression to systematically characterize smoking-stratified Th₁/Th₂ inflammatory phenotypes in pre-COPD high-risk individuals defined by the 2024 Chinese Primary Care COPD Guidelines. We identified distinct immune polarization patterns that may be undetectable by conventional mean-based statistics: smoking-related high-risk individuals exhibited Th₁-dominant immunity, while never-smoking counterparts showed Th₂-skewed polarization. These stage-specific, quantile-dependent biomarker changes offer preliminary insights into early immune dysregulation and provide a preliminary basis for the development of stratified screening strategies in primary care.

In the smoking subgroup, quantile regression revealed elevated neutrophil counts/percentage and reduced lymphocyte proportion in both high-risk individuals and COPD patients across most quantiles, reflecting early Th₁-skewed immunity. Cigarette smoke activates Th₁ innate immune responses, driving sustained neutrophil-mediated low-grade inflammation and impaired adaptive immune homeostasis^[16-18]. As core effector cells of Th₁-mediated inflammation, neutrophils are continuously recruited and activated under persistent smoke stimulation, driving sustained systemic low-grade inflammation. The declined lymphocyte percentage indicates imbalanced adaptive immune homeostasis caused by excessive Th₁ polarization^[18]. Our data support that Th₁-dominant inflammation emerges long before the development of irreversible airflow limitation, representing a potential

window for early intervention.

Notably, erythrocyte disturbances (reduced Hb, MCH, MCHC; elevated RDW-CV) were more prominent in high-risk individuals than in COPD patients, representing an inverse U-shaped trajectory that differs from the classic assumption of progressive deterioration with disease stage^[14]. This unexpected finding may be explained by early inflammatory consumption and oxidative stress that disrupt erythropoiesis in the preclinical phase, while compensatory erythropoiesis, comorbid polycythemia and disease adaptation mask these alterations in established COPD.

Th₁-related pro-inflammatory cytokines exhibited distinct stage-specific and quantile-dependent expression patterns. Consistent with previous evidence, COPD patients showed significantly higher IL-6 at $q = 0.50$ and IL-8 at $q = 0.25$ ^[4, 19, 20]. Notably, TNF- α was only elevated at lower quantiles in the high-risk group, with no significant difference between COPD patients and healthy controls. A notable finding was that TNF- α was significantly elevated only at $q=0.25$ in the high-risk group, with no difference between COPD patients and healthy controls. This quantile-restricted elevation may explain the inconsistent TNF- α findings in previous studies that focused exclusively on mean values. Two complementary mechanisms may underlie this observation: first, TNF- α acts as an early trigger of Th₁-mediated inflammation that is transiently upregulated in the preclinical phase, while persistent chronic inflammation in established COPD induces immune cell exhaustion and reduced TNF- α production^[17, 21]; second, endogenous anti-inflammatory pathways are progressively activated with disease progression, counteracting TNF- α elevation in COPD patients but not yet fully functional in pre-COPD individuals^[16].

In the never-smoking subgroup, distinct Th₂-skewed immune and hematological profiles were observed. EOS count and percentage were significantly higher in COPD patients across most quantiles, but not in high-risk individuals, indicating that overt eosinophilic inflammation develops only after persistent airflow limitation occurs^[5, 22, 23]. This finding challenges the common notion that eosinophilic inflammation is an early event in all COPD subtypes and suggests that it may represent a late-stage feature of never-smoking COPD. Notably, never-smoking high-risk individuals

exhibited significantly reduced MCHC across all three quantiles ($q=0.25$, 0.50 and 0.75), which was also lower than in COPD patients at $q=0.50$. This progressive, quantile-dependent reduction may reflect cumulative chronic low-grade inflammation that interferes with erythropoiesis, iron utilization and hemoglobin metabolism^[13, 14, 24]. The consistency of this finding across the biomarker distribution suggests that MCHC may be a promising early biomarker for never-smoking pre-COPD.

As a key Th₂ cytokine, IL-5 was significantly elevated only at $q=0.50$ in never-smoking pre-COPD individuals, with no difference between COPD patients and healthy controls. This stage-specific expression pattern suggests that Th₂ activation is an early, transient event in never-smoking pre-COPD, followed by progressive immune remodeling in established disease: as disease progresses, enhanced anti-inflammatory compensation and progressive Th₁ immune dysfunction suppress Th₂ activity, leading to normalized IL-5 levels^[5, 6]. IL-12p70 and IFN- γ were gradually reduced in high-risk and COPD groups, indicating progressive Th₁ immune dysfunction^[4, 6, 17, 20]. Collectively, never-smoking individuals display a unique pattern of early transient Th₂ activation and progressive Th₁ suppression.

Correlation analysis further explored the dose-dependent effect of smoking on inflammatory phenotypes. In the total study population, smoking index was positively correlated with neutrophil-related hematological indicators and Th₁-type pro-inflammatory cytokines, while inversely associated with lymphocyte percentage. Intriguingly, in the smoking subgroup, the significant correlations between smoking index and routine blood parameters were no longer observed, while the positive correlations with Th₁-related cytokines (IL-6 and TNF- α) remained. This dissociation suggests that cumulative smoking exposure may specifically amplify Th₁-type pro-inflammatory responses, whereas the observed hematological changes in the total population are primarily driven by the difference between smokers and never-smokers rather than by smoking dose itself. This finding provides preliminary evidence that smoking is a key driver of Th₁-dominant inflammation in COPD.

Several important limitations of this study should be acknowledged. First, this was an exploratory study with a relatively small sample size, which limited statistical power

to detect small effect sizes and rare associations. Larger studies are needed to confirm our preliminary findings. Second, as a single-center cross-sectional study, we can only establish associations rather than causal relationships between smoking status, Th₁/Th₂ phenotypes and biomarker changes. Large-scale multicenter prospective cohort studies are needed to validate whether the identified quantile-specific biomarkers can predict progression from pre-COPD to confirmed COPD and response to early intervention^[25]. Third, although we adjusted for age, sex and smoking index, residual confounding by unmeasured factors such as alcohol consumption, occupational exposure and comorbidities cannot be completely excluded^[5, 26, 27]. Fourth, we only measured peripheral blood biomarkers, and the correlation between systemic and airway-local Th₁/Th₂ phenotypes requires further validation using induced sputum samples.

Despite these limitations, our study has potential clinical implications for primary care COPD prevention. Routine blood parameters and inflammatory cytokines are low-cost, non-invasive and widely accessible, making them suitable for large-scale community screening. Our preliminary findings support the potential development of a stratified early screening approach: smoking-related pre-COPD individuals may benefit from targeted monitoring of low-quantile TNF- α , neutrophil count and RDW-CV; never-smoking high-risk individuals may require closer attention to mid-quantile IL-5 and MCHC. Quantile regression shows promise as a tool for identifying subtle, distribution-specific early abnormalities missed by conventional statistics. Future studies should focus on validating these biomarkers in independent populations, developing and validating quantile-specific cut-off values that can distinguish pre-COPD high-risk individuals from healthy controls, and evaluating their performance in real-world primary care screening settings.

Conclusions

This exploratory community-based cross-sectional study provides preliminary evidence that smoking status stratifies distinct Th₁/Th₂ inflammatory phenotypes in pre-COPD high-risk individuals: smoking-related individuals exhibit early Th₁-dominant immunity, while never-smoking counterparts show Th₂-skewed

polarization. Quantile regression identifies quantile-specific early immune alterations, including neutrophilia and isolated low-quantile TNF- α elevation in smokers, and mid-quantile IL-5 elevation in never-smokers. These low-cost peripheral blood markers may support stratified early screening of pre-COPD populations in primary care.

Declarations

Abbreviations

CI: Confidence interval

COPD: Chronic obstructive pulmonary disease

COPD-SQ: COPD Screening Questionnaire

EDTA: Ethylenediaminetetraacetic acid

EOS: Eosinophil

FEV₁: Forced expiratory volume in 1 second

FVC: Forced vital capacity

Hb: Hemoglobin

HCT: Hematocrit

IFN- α : Interferon- α

IFN- γ : Interferon- γ

IL: Interleukin

MAE: Mean absolute error

MCV: Mean corpuscular volume

MCH: Mean corpuscular hemoglobin

MCHC: Mean corpuscular hemoglobin concentration

MPV: Mean platelet volume

NSAIDs: Non-steroidal anti-inflammatory drugs

PCT: Plateletcrit

PDW: Platelet distribution width

RDW-CV: Red blood cell distribution width-coefficient of variation

RDW-SD: Red blood cell distribution width–standard deviation

Th₁: T helper type 1

Th₂: T helper type 2

TNF- α : Tumor necrosis factor- α

Ethics approval and consent to participate

The study protocol was reviewed and approved by the Ethics Committee of Shanghai Fifth People's Hospital, Fudan University (Approval No. 2023-153). Written informed consent was obtained from all participants prior to any study-related procedures.

Consent for publication

All participants have provided written informed consent for the use of their anonymized research data in academic publications. No individually identifiable patient information is contained in this manuscript.

Availability of data and materials

The datasets generated and analyzed during this study are not publicly available due to privacy and ethical restrictions on participant data. Reasonable requests for data access can be directed to the corresponding authors.

Competing Interests

All authors declare no competing interests. The funders had no involvement in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Funding

This work was supported by the Minhang District Natural Science Research Project (No. 2023MHZ059, No. 2024MHZ059) and the Minhang District Medical Characteristic Specialty Construction Project (No. 2025MWFC07).

Authors' contributions

Xiaohua Zhang and Juan Xie contributed equally to this work as co-first authors. X.Z.: Data curation, Formal analysis, Methodology, Writing – original draft. J.X.: Methodology, Writing – review & editing. S.Y.: Investigation, Participant recruitment. B.W.: Resources, Laboratory measurements. J.S.: Conceptualization, Methodology, Writing – review & editing (Corresponding Author). X.H.: Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration (Corresponding Author). All authors have read and approved the final version of the

manuscript.

Acknowledgements

The authors thank all residents who participated in this study, as well as the primary care staff of Jiangchuan Community Health Service Center for their support in participant recruitment and data collection.

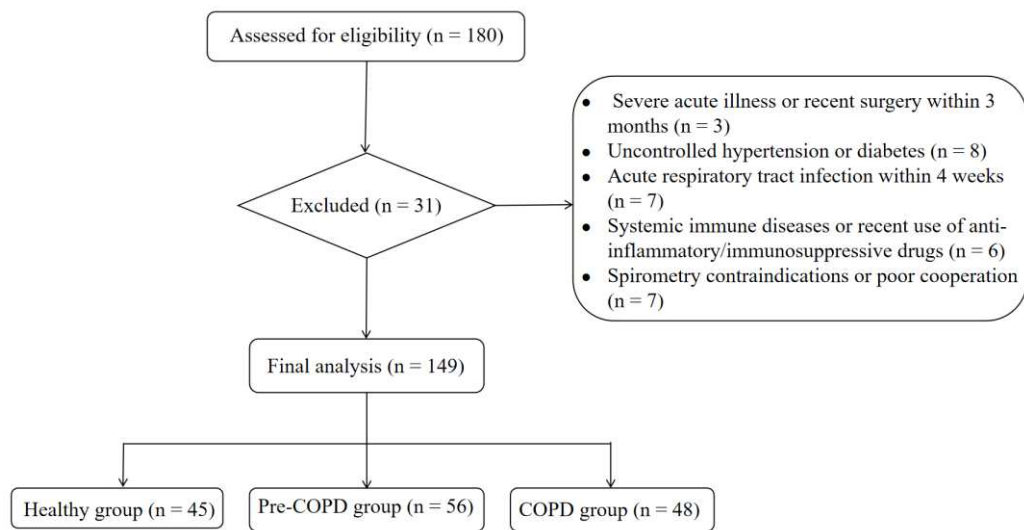


Fig. 1 Flow chart of participant screening and grouping.

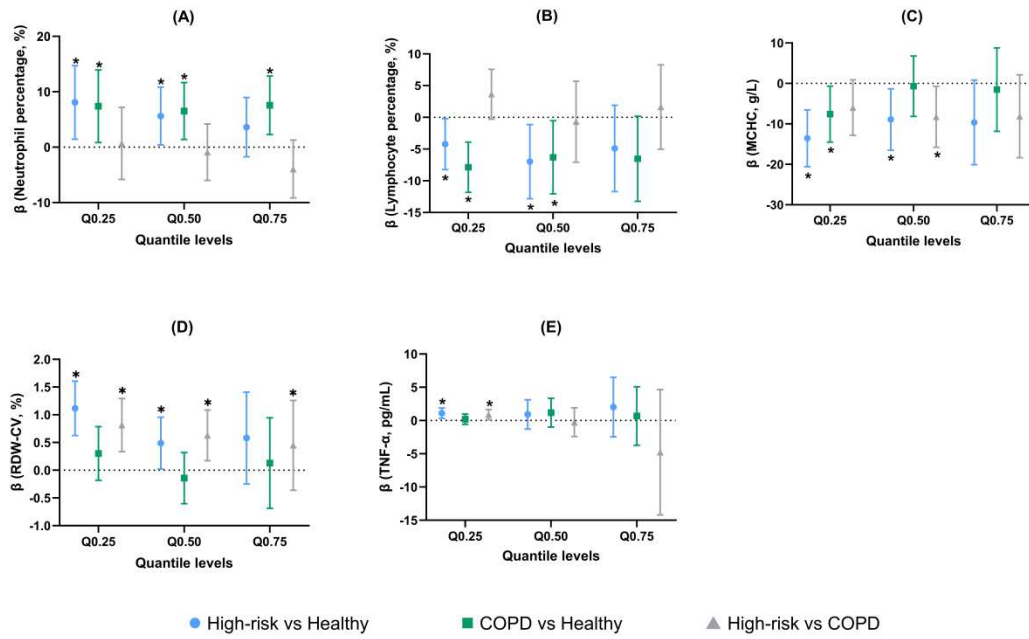


Fig. 2 Quantile regression analysis of key biomarkers in the smoking subgroup. Panels (A–E) show adjusted regression coefficients (β) with 95% confidence intervals at $q = 0.25, 0.50$ and 0.75 : (A) neutrophil percentage, (B) lymphocyte percentage, (C) mean corpuscular hemoglobin concentration (MCHC), (D) red blood cell distribution width-coefficient of variation (RDW-CV), (E) tumor necrosis factor- α (TNF- α). All models were adjusted for age, sex and smoking index. The dashed horizontal line at $\beta = 0$ indicates no between-group difference. * $P < 0.05$ indicates statistically significant between-group differences.

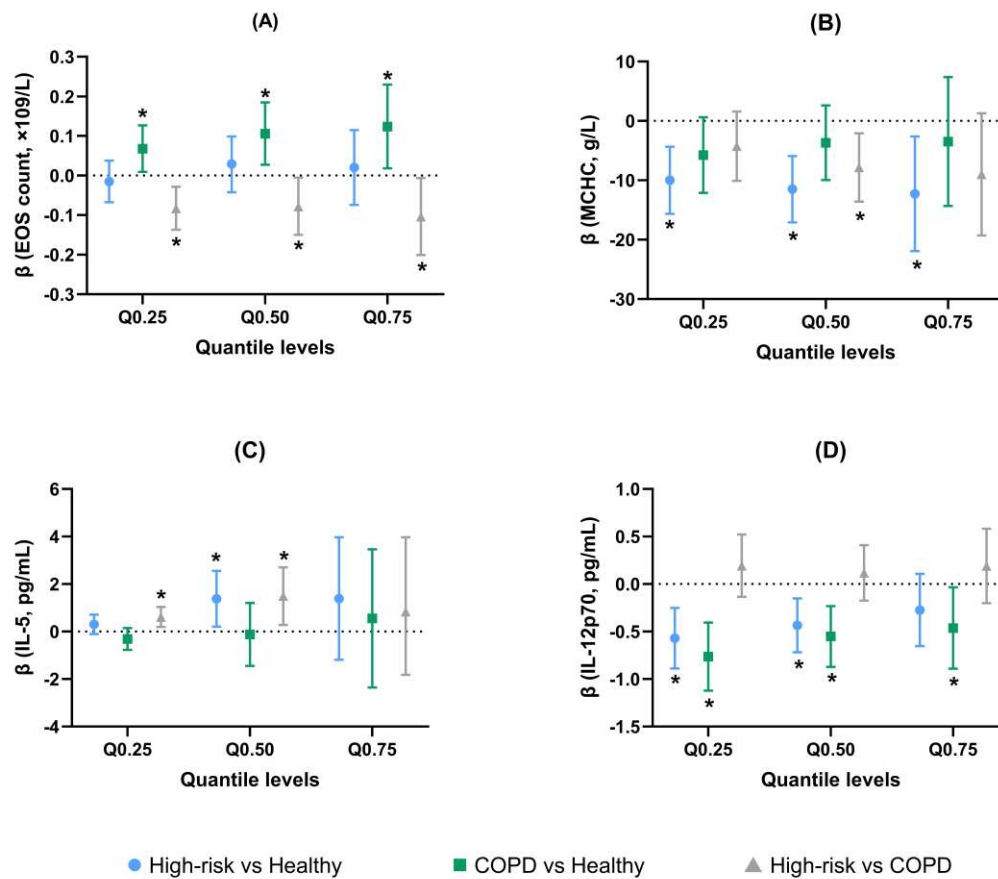


Fig. 3 Quantile regression analysis of key biomarkers in the never-smoking subgroup. Panels (A–D) show adjusted regression coefficients (β) with 95% confidence intervals at $q = 0.25, 0.50$ and 0.75 : (A) eosinophil (EOS) count, (B) mean corpuscular hemoglobin concentration (MCHC), (C) interleukin-5 (IL-5), and (D) interleukin-12p70 (IL-12p70). All models were adjusted for age and sex. The dashed horizontal line at $\beta=0$ indicates no between-group difference. * $P < 0.05$ indicates statistically significant between-group differences.

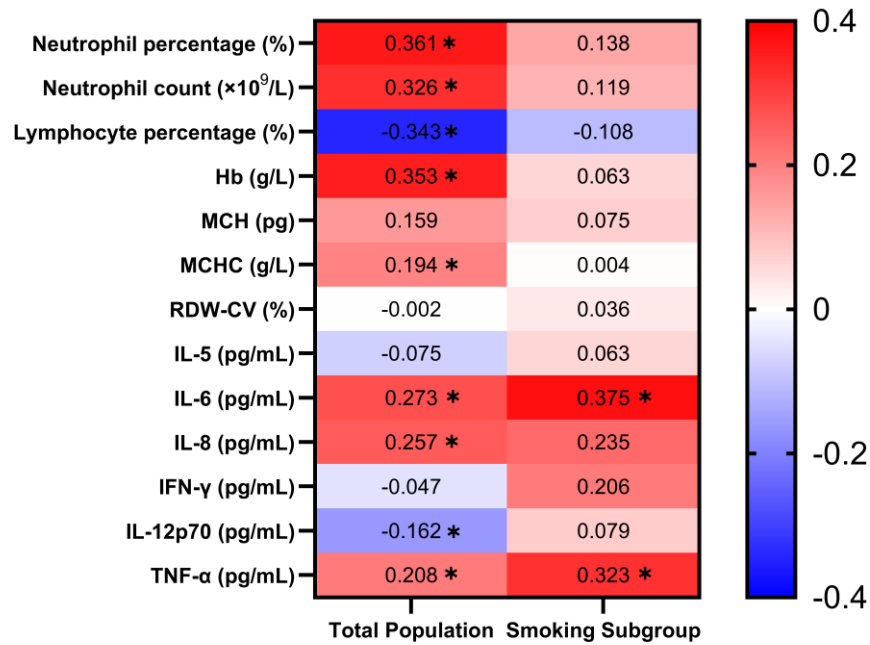


Fig. 4 Spearman's rank correlation between smoking index and core biomarkers in the total population and smoking subgroup. The heatmap shows Spearman's correlation coefficients (r_s) between smoking index and key hematological and inflammatory biomarkers, stratified by total population (left) and smoking subgroup (right). Red and blue represent positive and negative correlations, respectively (color scale: -0.4 to 0.4). Asterisks (*) denote statistically significant correlations ($P < 0.05$).

References

- [1] Safiri S, Carson-Chahhoud K, Noori M, et al. Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: results from the Global Burden of Disease Study 2019[J]. *BMJ*, 2022,378:e69679.
- [2] Wang C, Xu J, Yang L, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study[J]. *Lancet*, 2018,391(10131):1706-1717.
- [3] Christenson S A, Smith B M, Bafadhel M, et al. Chronic obstructive pulmonary disease[J]. *Lancet*, 2022,399(10342):2227-2242.
- [4] Barnes P J. Inflammatory mechanisms in patients with chronic obstructive pulmonary disease[J]. *The Journal of allergy and clinical immunology*, 2016,138(1):16-27.
- [5] GOLD G I F C. Global Strategy for the Prevention, Diagnosis and Management of COPD: 2024 Report[EB/OL]. [2025-09-05]. <https://goldcopd.org/2024-gold-report/>.
- [6] Barnes P J. Inflammatory endotypes in COPD[J]. *Allergy*, 2019,74(7):1249-1256.
- [7] Barnes P J. Cellular and molecular mechanisms of asthma and COPD[J]. *Clin Sci (Lond)*, 2017,131(13):1541-1558.
- [8] Murphy D E, Panos R J. Diagnosis of COPD and clinical course in patients with unrecognized airflow limitation[J]. *Int J Chron Obstruct Pulmon Dis*, 2013,8:199-208.
- [9] Zhou Y M, Chen S Y, Tian J, et al. Development and validation of a chronic obstructive pulmonary disease screening questionnaire in China[J]. *Int J Tuberc Lung Dis*, 2013,17(12):1645-1651.
- [10] Shen X, Yang H, Lan C, et al. Screening performance of COPD-PS scale, COPD-SQ scale, peak expiratory flow, and their combinations for chronic obstructive pulmonary disease in the primary healthcare in Haicang District, Xiamen City[J]. *Frontiers in medicine*, 2024,11:1357077.
- [11] Chinese Medical Association, Chinese Medical Association Journal House, Chinese Medical Association General Practice Branch, et al. Chinese Guidelines for Primary Care Diagnosis and Management of Chronic Obstructive Pulmonary Disease (2024)[J]. *Chinese Journal of General Practitioners*, 2024,23(6):578-602.
- [12] Hoenderdos K, Condcliffe A. The neutrophil in chronic obstructive pulmonary disease[J]. *Am J Respir Cell Mol Biol*, 2013,48(5):531-539.
- [13] Tariq S, Ismail D, Thapa M, et al. Chronic Obstructive Pulmonary Disease and Its Effect on Red Blood Cell Indices[J]. *Cureus*, 2023,15(3):e36100.
- [14] Liu S, Zhang H, Zhu P, et al. Predictive role of red blood cell distribution width and hemoglobin-to-red blood cell distribution width ratio for mortality in patients with COPD: evidence from NHANES 1999-2018[J]. *BMC pulmonary medicine*, 2024,24(1):413.
- [15] Young R P, Hopkins R J. Interleukin-6 and statin therapy: potential role in the management of COPD[J]. *Respiratory research*, 2013,14(1):74.
- [16] Celli B R, Christenson S A, Rabe K F, et al. Current Smoker: A Clinical Chronic Obstructive Pulmonary Disease Phenotype Affecting Disease Progression and Response to Therapy[J]. *American journal of respiratory and critical care medicine*, 2025,211(5):729-736.
- [17] Li H, Wang Y, Duan H, et al. Immune cell regulatory networks in chronic obstructive pulmonary disease: mechanistic analysis from innate to adaptive immunity[J]. *Frontiers in immunology*, 2025,16:1651808.

- [18] Hoesterey D T, Dang H, Markovic D, et al. Neutrophil-to-Lymphocyte Ratio as a Biomarker in Clinically Stable Chronic Obstructive Pulmonary Disease: SPIROMICS Cohort[J]. *Annals of the American Thoracic Society*, 2025,22(12):1881-1890.
- [19] Tanni S E, Pelegrino N R, Angeleli A Y, et al. Smoking status and tumor necrosis factor-alpha mediated systemic inflammation in COPD patients[J]. *Journal of inflammation (London, England)*, 2010,7:29.
- [20] Chen G, Mu Q, Meng Z. Cigarette Smoking Contributes to Th1/Th2 Cell Dysfunction via the Cytokine Milieu in Chronic Obstructive Pulmonary Disease[J]. *International journal of chronic obstructive pulmonary disease*, 2023,18:2027-2038.
- [21] Wu B, Li G, Hu X, et al. Threshold-Triggered Immune Dysregulation Precedes Clinical COPD: A Stage-Specific Diagnostic Model Combining Cytokine Profiling and Lymphocyte Phenotyping[J]. *International journal of general medicine*, 2025,18:6899-6908.
- [22] David B, Bafadhel M, Koenderman L, et al. Eosinophilic inflammation in COPD: from an inflammatory marker to a treatable trait[J]. *Thorax*, 2020,76(2):188-195.
- [23] Salvi S S, Brashier B B, Londhe J, et al. Phenotypic comparison between smoking and non-smoking chronic obstructive pulmonary disease[J]. *Respiratory research*, 2020,21(1):50.
- [24] Minegishi Y, Sato K, Nawa S, et al. Lower mean corpuscular hemoglobin levels as a predictive factor of future exacerbations in patients with chronic obstructive pulmonary disease[J]. *Respiratory Investigation*, 2025,63(1):183-190.
- [25] Qiao L, Wu X, Chen G, et al. Incidence and Risk Factors for Progression from Pre-COPD to COPD: A Systematic Review and Meta-analysis[J]. *International journal of chronic obstructive pulmonary disease*, 2026,21:566518.
- [26] Bongers Q, Comellas A P. Impact of occupational exposures in patients with chronic obstructive pulmonary disease: current understanding and knowledge gaps[J]. *Current opinion in pulmonary medicine*, 2025,31(2):98-105.
- [27] Ruan H, Lei S, Zhao H, et al. Association between comorbidity and chronic obstructive pulmonary disease: a systematic review and meta-analysis of Mendelian randomization studies[J]. *Therapeutic advances in respiratory disease*, 2025,19:1591039575.