Potential pre-COPD indicators in association with COPD development and COPD prediction models in Chinese: a prospective cohort study

Jing Fan,[a,b] Liwen Fang,[a,*,c] Shu Cong,[a] Yang Zhang,[a] Xiao Jiang,[a] Ning Wang,[a] and Yahong Chen[b,**,c]

[a]National Center for Chronic and Non-communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, 27 Nanwei Road, Xicheng District, Beijing 100050, China
[b]Department of Pulmonary and Critical Care Medicine, Peking University Third Hospital, 49 Huayuan North Road, Haidian District, Beijing 100191, China

The Lancet Regional Health - Western Pacific 2024;44: 100984

Published Online xxx
https://doi.org/10.1016/j.lanwpc.2023.100984

Summary

Background Lung injury might take place before chronic obstructive pulmonary disease (COPD) occurs. A clearer definition of "pre-COPD" based on the effects of potential indicators on increasing risk of COPD development and a prediction model involving them are lacking.

Methods A total of 3526 Chinese residents without COPD aged 40 years or older derived from the national cross-sectional survey of COPD surveillance in 2014–2015 were followed up for a mean of 3.59 years. We examined the associations of chronic bronchitis, preserved ratio impaired spirometry (PRISm), low peak expiratory flow (PEF), spirometric small airway dysfunction (sSAD), low maximal mid-expiratory flow (MMEF), low forced expiratory flow 50% of pulmonary volume (FEF50), and low FEF75 with subsequent COPD and constructed a prediction model with LASSO-Cox regression.

Findings 235 subjects in the cohort developed COPD during the follow-up. Subjects with PRISm, low PEF, sSAD, low MMEF, low FEF50, and low FEF75 had an increased risk of developing COPD (adjusted hazard ratio [HR] ranging from 1.57 to 3.01). Only chronic bronchitis (HR 2.84 [95% CI 1.38–5.84] and 2.94 [1.43–6.04]) and sSAD/low MMEF (HR 2.74 [2.07–3.61] and 2.38 [1.65–3.43]) showed effects independent of the other indicators and their concurrence had the strongest effect (HR 5.89 and 4.80). The prediction model including age, sex, low MMEF, low FEF50, and indoor exposure to biomass had good performance both internally and temporally. The corrected C-index was 0.77 (0.72–0.81) for discrimination in internal validation. For temporal validation, the area under the receiver operating characteristic curve was 0.73 (0.63–0.83). Good calibration was indicated in plot for internal validation and by Hosmer–Lemeshow test for temporal validation.

Interpretation Individuals with concurrent chronic bronchitis and sSAD/low MMEF indicating pre-COPD optimally require more high attention from physicians. Our prediction model could serve as a multi-dimension tool to predict COPD comprehensively.

Funding The Ministry of Finance and the Ministry of Science and Technology of the People's Republic of China and the National Natural Science Foundation of China.

摘要

背景 在慢性阻塞性肺疾病(chronic obstructive pulmonary disease, COPD)出现前, 肺损伤可能已经发生。目前尚缺乏依据潜在指标对升高COPD发生风险的影响而产生的更明确的"慢阻肺前期"的定义， 以及包含这些潜在指标的慢阻肺发病预测模型。

方法 本研究基于2014–2015年中国居民慢阻肺监测横断面调查，对3526名40岁及以上的非慢阻肺居民进行了平均3.59年的随访；旨在分析慢性支气管炎、保留比值的肺功能损伤 (preserved ratio impaired spirometry, PRISm)、低呼气峰流速 (peak expiratory flow, PEF)、小气道功能障碍 (spirometric small airway dysfunction, sSAD)、低最大呼气中期流速 (maximal mid-expiratory flow, MMEF)、低呼出50%用力肺活量时的最大呼气流速

*Corresponding author.
**Corresponding author.
E-mail addresses: fangliwen@ncncd.chinacdc.cn (L. Fang), chenyahong@vip.sina.com (Y. Chen).
[c]Joint corresponding authors.

(forced expiratory flow 50% of pulmonary volume, FEF50)、低FEF75与COPD发生的关系，并采用LASSO-Cox回归构建预测模型。

结果 本前瞻性队列中共有235名研究对象在随访期间发展为COPD患者。PRISm、低PEF、sSAD、低MMEF、低FEF50和低FEF75者发生COPD的风险增加(调整后的风险比 [hazard ratio, HR] 在1.57至3.01之间)。只有慢性支气管炎 (HR 2.84 [95% CI 1.38–5.84] 和2.94 [1.43–6.04]) 和sSAD/低MMEF (HR 2.74 [2.07–3.61] 和2.38 [1.65–3.43]) 的效应独立于其他指标，且慢性支气管炎和sSAD共存或慢性支气管炎和低MMEF共存的效应最强 (HR 5.89和4.80)。本研究构建的预测模型因子包括年龄、性别、低MMEF、低FEF50和室内生物燃料暴露。该模型在内部验证和外部验证中均表现出良好的预测性能。内部验证中该模型的区分度即校正后C-index为0.77 (0.72–0.81)。外部验证中受试者工作特征曲线下面积为0.73 (0.63–0.83)。该模型在内部验证的校准图和外部验证的Hosmer-Lemeshow检验中均显示出良好的校准度。

解读 慢性支气管炎和sSAD/低MMEF是提示COPD前期的最佳指标，同时存在这两种表现的个体需引起临床医生的高度重视。本研究构建的预测模型可以作为一个多维度的综合预测COPD的工具。

资助 中华人民共和国财政部、科学技术部和国家自然科学基金。

Editor note: This translation in Chinese was submitted by the authors and we reproduce it as supplied. It has not been peer reviewed. Our editorial processes have only been applied to the original abstract in English, which should serve as reference for this manuscript.

Keywords: COPD; Pre-COPD; Prediction model

Introduction

Chronic obstructive pulmonary disease (COPD) is one of the leading causes of death in the world.[1] According to the Global Initiative for Chronic Obstructive Lung Disease, the gold standard for the definition of COPD is post-bronchodilator forced expiratory volume in 1 s (FEV_1)/forced vital capacity (FVC) less than 70%.[2] However, it takes a long period of time to develop COPD, which might be induced by smoking, indoor air pollution, occupational exposure, and other environmental factors.[2] Since lung injury might take place before FEV_1/FVC becomes less than 70%, it is significant to define a range of clinical indicators comprising the term "pre-COPD" to identify individuals with normal spirometry at present but at increased risk of developing COPD in the future.[3]

Some indicators have been proposed by previous studies to imply "pre-COPD" potentially, such as symptoms (e.g., chronic bronchitis), functional (e.g., small airway dysfunction), or imaging abnormalities.[3–10] However, each of these studies on progression to COPD was focused on only one sole indicator to identify individuals at higher risk of developing COPD. Besides, some were only limited in specific populations rather than the general population and Chinese population was not involved in any of them.[4–10] Hence, the distinction and the priority of these indicators in increasing the risk of developing COPD and their predictive values remain unknown. A clearer, refined and clinically operable definition of "pre-COPD" and a prediction model involving "pre-COPD" which might be defined from symptom, function, and structure are lacking for earlier detection of COPD.[11,12]

Using data from the national cross-sectional study of COPD surveillance in China and the follow-up of the subjects in certain sites, we sought to (1) examine the associations of potential pre-COPD indicators with COPD development to develop a clear cognition about pre-COPD, and (2) develop and validate a prediction model for COPD development involving potential indicators for pre-COPD and risk factors for COPD.

Methods

Study design and participants

We did a prospective cohort study in subjects derived from the COPD surveillance in China. The COPD surveillance is a national cross-sectional survey conducted every 5 years and covered a representative sample of Chinese population aged 40 years or older. Using a complex, stratified, multistage cluster sampling strategy, 66,752 individuals were sampled and investigated with questionnaires and acceptable post-bronchodilator spirometry tests in 125 sites (counties/districts) from 31 provinces in China in 2014–2015.[13] After that fourteen sites intended for follow-up were selected across six regions of China–North, Northeast, East, Central, South, and Southwest China, where 7576 subjects had been investigated with acceptable post-bronchodilator spirometry examinations in 2014–2015. During 2018–2020, a total of 5306 subjects of them, 4419 of whom had no

Research in context

Evidence before this study

We searched PubMed, Wanfang Database and China National Knowledge Infrastructure for articles published up to December, 2022, using terms "chronic obstructive pulmonary disease/COPD", "chronic bronchitis", "cough", "sputum production", "phlegm", "chronic mucus hypersecretion", "preserved ratio impaired spirometry/PRISm", "peak expiratory flow/PEF", "small airway function", "small airway dysfunction", "small airway disease", "emphysema" and "pre-COPD" in English and Chinese. Some indicators have been proposed by previous studies to imply "pre-COPD" potentially from symptoms, functional or imaging abnormalities (Chronic cough/phlegm: incidence rate ratio 1.85 (95% CI, 1.17–2.93); chronic bronchitis: hazard ratio 1.37 (0.98–1.92); PRISm: hazard ratio 2.48 and odds ratio 3.75; FEV_3/FEV_6 < LLN: hazard ratio 2.11; N_2-slope increase in quintiles: hazard ratio 1.63; emphysema: odds ratio 4.38 and hazard ratio 5.14). However, each of these studies on progression to COPD was focused on only one sole indicator to identify individuals at higher risk of developing COPD. Besides, some of them were only limited in specific populations rather than the general population and Chinese population was not involved in any of them. Therefore, the distinction of these indicators in increasing the risk of developing COPD and their predictive values remains uncertain and a clearer, refined and clinically operable definition of "pre-COPD" and a prediction model involving "pre-COPD" which could be defined from symptom, function, and structure are lacking for earlier detection of COPD.

Added value of this study

To our knowledge, this study is the first one to provide prospective evidence for effects of potential pre-COPD indicators on predicting progression to COPD in one Chinese cohort. We found that subjects with PRISm, low PEF, spirometric small airway dysfunction (sSAD), low MMEF, low FEF50, and low FEF75 had an increased risk of developing COPD (hazard ratio 1.57, 1.65, 2.77, 3.01, 2.52, and 2.34) after adjusting for other risk factors. Only chronic bronchitis and sSAD/low MMEF showed independent effects when adjusted for the other indicators and their concurrence had the strongest effect (hazard ratio 5.89 and 4.80) compared to each of them alone. The prediction model for COPD development including age, sex, low MMEF, low FEF50, and indoor exposure to biomass had good performance both internally and temporally. The C-index and corrected C-index was 0.7704 and 0.7660 for discrimination in internal validation. For temporal validation, the area under the receiver operating characteristic curve was 0.730. The cut-off value of the prognostic index was 3.630 to discriminate individuals at a high or low risk of COPD, at which the sensitivity and the specificity was 0.633 and 0.774 respectively and the positive and negative predictive value was 0.165 and 0.968.

Implications of all the available evidence

The presence of chronic bronchitis and sSAD/low MMEF predicted COPD independently and were the most optimal of those investigated in Chinese population. Physicians should pay more attention to patients with chronic bronchitis and sSAD/low MMEF concurrently. Our prediction model involving these pre-COPD indicators and other risk factors may serve as a multi-dimension tool to predict COPD comprehensively. Taking advantage of pre-COPD and the prediction model, we are able to identify the high risk individuals for COPD and initiate interventions at an earlier stage. Further studies are needed to illustrate the particular risks of COPD and develop specific prediction models in populations with or without exposures to specific COPD risk factors to understand the natural history of COPD in various contexts and facilitate interventions and treatments accordingly.

airflow limitation in 2014–2015, were initially contacted for the survey at the follow-up. Among the 4419 subjects without COPD, 4093 subjects participated in the interview at the follow-up, 3526 of whom had acceptable post-bronchodilator spirometry examinations then, 47 subjects were found dead before the contact and 279 were lost to follow-up. Finally, 3526 subjects were included in the cohort for the association analysis and used as the training set for the prediction model (Fig. 1a).

An entry date to the cohort was defined for each subject as the date of the first survey in 2014–2015. An exit date was defined as the date of the follow-up survey during 2018–2020.

A total of 482 subjects took part in both the first (2014–2015) and second (2019–2020) surveys of COPD surveillance with acceptable post-bronchodilator spirometry examinations and had no airflow limitation in 2014–2015. In order to ensure the subjects in the training and validation datasets were not overlapped, 21 subjects living in the fourteen sites for follow-up were excluded and the remaining 461 subjects were used as the validation cohort for the prediction model (Fig. 1b). The sites in the validation cohort covered all the seven regions over the country including Northwest China.

The study was approved by the ethics review committee of the National Center for Chronic and Non-Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention. Written informed consent was obtained from all participants.

Procedures

In the national cross-sectional surveys of COPD surveillance and the survey at follow-up, each participant was invited to an interview administered by local trained

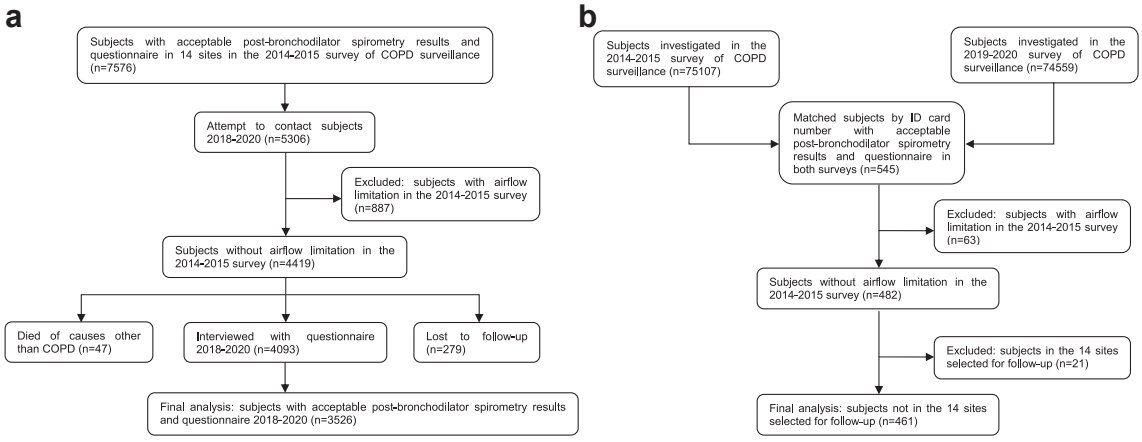

Fig. 1: Flowchart of study participants. a) the 3526 study participants in the training cohort; b) the 461 study participants in the validation cohort.

staff.[13] The questionnaire collected information on demographic characteristics, medical history, respiratory symptoms and exposure to risk factors for COPD. The definitions of smoking status, indoor exposure to biomass, indoor exposure to coal, and exposure to dust or chemicals in the workplace were described previously[13] and provided in the appendix. History of emphysema, coronary heart disease and hypertension was defined as the self-reported disease diagnosed by township or higher level hospitals before subjects were investigated in the first survey.

Pre- and post-bronchodilator spirometry were performed for eligible participants by trained staff using the same brand of spirometer and following the same procedure to measure FVC, FEV_1, peak expiratory flow (PEF), maximal mid-expiratory flow (MMEF), and forced expiratory flow 50% and 75% of pulmonary volume (FEF50 and FEF75).[13] Quality control of spirometry was described previously[13] and provided in the Appendix.

The primary outcome of this study is the occurrence of COPD defined as post-bronchodilator FEV_1/FVC <70%[2] in the survey at the follow up or in the second survey of COPD surveillance.

All seven potential indicators (chronic bronchitis, preserved ratio impaired spirometry (PRISm), low PEF, spirometric small airway dysfunction (sSAD), low MMEF, low FEF50, and low FEF75) were defined under the condition of post-bronchodilator FEV_1/FVC ≥70% in the survey in 2014–2015. Chronic bronchitis were defined as chronic cough for more than four times per day, 4 days or more per week, and at least 3 months in each of two consecutive years and chronic sputum production for more than twice per day, 4 days or more per week, and at least 3 months in each of two consecutive years.[14] Since an increase in FEV_1 ≥12% and ≥200 mL from the pre-bronchodilator value or a change >10% of its predicted value indicates a positive bronchodilator response,[15,16] PRISm was defined as post-bronchodilator FEV_1 <80% predicted to minimize the potential influence caused by the airway reversibility in some individuals.[17] Low PEF was defined as pre-bronchodilator PEF <80% predicted.[18] According to results of previous studies and available indicators for small airway function in our examinations, sSAD was defined as at least two of the three indicators of pre-bronchodilator MMEF, FEF50, and FEF75 less than 65% predicted.[19,20] In order to explore the priority of these three indicators in predicting COPD, we similarly defined low MMEF, low FEF50 and low FEF75 as its pre-bronchodilator value less than the 65% predicted respectively. Since changes in forced expiratory flows (such as PEF or MMEF) are highly variable and significantly affected by changes in FVC, pre- and post-bronchodilator measurements are not comparable.[15] Hence, pre-bronchodilator PEF, MMEF, FEF50, and FEF75 were used in our study. Predicted values of FEV_1, PEF, and MMEF for normal lung function were obtained from a national study in Chinese.[21] Predicted values of FEF50 and FEF75 were obtained from an official statement of the European Respiratory Society.[22]

Statistical analysis

The cumulative incidence of COPD was calculated and compared by univariate Cox proportional hazard regression among subjects with different characteristics in the training cohort.

The relationships between potential indicators/their combinations and COPD development

The crude hazard ratios (HRs) of COPD were estimated for potential indicators which were significant in the univariate Cox model and their adjusted HRs were estimated in respective multivariable models with covariates including baseline sex, age, smoking status,

exposure to dust or chemicals in the workplace, indoor exposure to coal, and indoor exposure to biomass (model 1). Smoking and the environmental covariates which were generally accepted risk factors for COPD were considered because they were modifiable factors and could be improved in the lifetime. To consider these potential indicators all at once, only indicators which had been significant in the univariate Cox model were included in model 2 and 3. Chronic bronchitis, PRISm, low PEF, and sSAD were included in model 2 while low MMEF, low FEF50, and low FEF75 substituted for sSAD in model 3. Combinations of significant indicators in model 2 and 3 were modeled to estimate their HRs with history of coronary heart disease and hypertension adjusted for additionally. Collinearity was tested for all the variables in the models but none were found. The associations between any two of the indicators at baseline were tested by chi-square.

For sensitivity analysis, we examined if the effect of each potential indicator on COPD development still existed after excluding individuals with every other indicator respectively. The purpose of this was to verify if the associations between indicators and COPD development were consistent in various populations or they were conditional or dependent on the other indicators.

The prediction model for COPD development

Least absolute shrinkage and selection operator (LASSO) regression model[23,24] was used to screen for predictors involved in the final prediction model developed by Cox regression. LASSO is one method of the regularization in which only one model is needed to be fit for each tuning parameter contained in the contraction penalty term, so the computational efficiency will be greatly improved. Second, it can properly select appropriate value for tuning parameter to optimize the deviation/variance of the residual sum of squares tradeoff to improve the fitness of the regression model. Third, it can be used to mitigate the over-fitting problem and can reduce the regression coefficient to zero increasing the interpretability of the model.[24] We used the minimum standard I-Standard Error method to determine the final number of predictors which was identified at the one standard error on the right from the minimum mean square error indicating good fitness of the model with the fewest predictors.[24] For internal validation, Harrell's C (C-Index) and Somers Dxy for censored data were calculated for discrimination evaluation and a C-Index >0.7 indicated good discrimination.[24] Optimism-corrected estimates of C-Index and Somers Dxy were provided using a bootstrap resampling process. Akaike information criterion (AIC)[25,26] indicating the goodness of fit of the model was estimated as well and calibration in plot predicting the probability of developing COPD at 5 years versus observed probability was evaluated through bootstrap resampling method.[24] For temporal validation, the prognostic index indicating the risk of developing COPD in future was calculated for each subject in the validation dataset as $\sum \beta_i \times X_i$ where β_i were the regression coefficients of covariates X_i from the prediction model and was used to determine the performance of the model with logistic regression. Sensitivity, that is the ability of the model to find the cases, was calculated as the number of actual incident COPD cases classified as at high risk of COPD by the model divided by the total number of actual incident COPD cases.[27] Specificity, that is the ability of the model to find the non-cases, was calculated as the number of actual non-incident COPD cases classified as at low risk of COPD by the model divided by the total number of actual non-incident COPD cases.[27] The cut-off value of prognostic index maximizing the sum of sensitivity and specificity to predict COPD development were estimated for the optimal point to discriminate between individuals at high and low risk of COPD. The area under the receiver operating characteristic (ROC) curve (AUC)[27] was calculated to evaluate the model's ability to discriminate between individuals at high and low risk of COPD in temporal validation. Youden's index defined as sensitivity – false positive error fraction maximizing the vertical line between ROC curve and diagonal line was calculated as sensitivity+specificity – 1, which reflected the overall ability of the model to predict subsequent COPD.[27] Positive predictive value was calculated as the probability of developing COPD actually among individuals classified as at high risk of COPD[27] and negative predictive value was calculated as the probability of not developing COPD actually among individuals classified as at low risk of COPD,[27] which are used by doctors to make decision when the test results are known. Hosmer–Lemeshow statistic (statistic C)[28] was calculated to evaluate the calibration degree in temporal validation which was the ability of the model to match the predictions to the actual disease outcomes.[24]

Statistical analyses were performed using Statistical Analysis System (SAS) version 9.4 and R 4.1.3 software. All tests were two sided. A P value less than 0.05 was considered to indicate statistical significance.

Role of the funding source

The funder of the study had no role in study design, data collection, data Formal analysis, data interpretation, writing of the report, or the decision to submit this report for publication.

Results

Table 1 presents the baseline characteristics and cumulative incidence of COPD (overall 6.66% [95% CI 5.84–7.49]) among 3526 individuals followed up in the cohort, out of whom 235 were found with an FEV_1/FVC <70% in the follow-up survey. With the average follow-up time 3.59 years the incidence density was

	Total[a]	Cases[a]	Non-cases[a]	Incidence (%, 95% CI)	P[b]
Overall	3526	235	3291	6.66 (5.84–7.49)	
Age (years)	54.5 (9.3)	59.7 (9.5)	54.1 (9.2)		<0.0001
Sex					<0.0001
Male	1487/3526 (42.2)	155/235 (66.0)	1332/3291 (40.5)	10.42 (8.87–11.98)	
Female	2039/3526 (57.8)	80/235 (34.0)	1959/3291 (59.5)	3.92 (3.08–4.77)	
Chronic bronchitis[c]					0.0039
No	3473/3524 (98.6)	227/235 (96.6)	3246/3289 (98.7)	6.54 (5.71–7.36)	
Yes	51/3524 (1.5)	8/235 (3.4)	43/3289 (1.3)	15.69 (5.71–25.67)	
PRISm					0.0206
No	3283/3526 (93.1)	210/235 (89.4)	3073/3291 (93.4)	6.40 (5.56–7.23)	
Yes	243/3526 (6.9)	25/235 (10.6)	218/3291 (6.6)	10.29 (6.47–14.11)	
Low PEF[c]					0.0001
No	2780/3483 (79.8)	163/231 (70.6)	2617/3252 (80.5)	5.86 (4.99–6.74)	
Yes	703/3483 (20.2)	68/231 (29.4)	635/3252 (19.5)	9.67 (7.49–11.86)	
sSAD[c]					<0.0001
No	2458/3483 (70.6)	109/231 (47.2)	2349/3252 (72.2)	4.43 (3.62–5.25)	
Yes	1025/3483 (29.4)	122/231 (52.8)	903/3252 (27.8)	11.90 (9.92–13.88)	
Low MMEF[c]					<0.0001
No	2878/3483 (82.6)	141/231 (61.0)	2737/3252 (84.2)	4.90 (4.11–5.69)	
Yes	605/3483 (17.4)	90/231 (39.0)	515/3252 (15.8)	14.88 (12.04–17.71)	
Low FEF50[c]					<0.0001
No	2488/3483 (71.4)	117/231 (50.7)	2371/3252 (72.9)	4.70 (3.87–5.53)	
Yes	995/3483 (28.6)	114/231 (49.4)	881/3252 (27.1)	11.46 (9.48–13.44)	
Low FEF75[c]					<0.0001
No	1209/3483 (34.7)	49/231 (21.2)	1160/3252 (35.7)	4.05 (2.94–5.16)	
Yes	2274/3483 (65.3)	182/231 (78.8)	2092/3252 (64.3)	8.00 (6.89–9.12)	
Emphysema[c]					0.9678
No	3510/3524 (99.6)	235/235 (100.0)	3275/3289 (99.6)	6.70 (5.87–7.52)	
Yes	14/3524 (0.4)	0/235 (0)	14/3289 (0.4)	0 (0–0)	
Smoking status[c]					<0.0001
Never smoker	2337/3521 (66.4)	115/235 (48.9)	2222/3286 (67.6)	4.92 (4.04–5.80)	
Former smoker	296/3521 (8.4)	21/235 (8.9)	275/3286 (8.4)	7.09 (4.17–10.02)	
Current smoker	888/3521 (25.2)	99/235 (42.1)	789/3286 (24.0)	11.15 (9.08–13.22)	
Exposure to dust or chemicals in the workplace[c]					0.5937
No	1743/3524 (49.5)	128/235 (54.5)	1615/3289 (49.1)	7.34 (6.12–8.57)	
Yes	1781/3524 (50.5)	107/235 (45.5)	1674/3289 (50.9)	6.01 (4.90–7.11)	
Indoor exposure to coal[c]					0.0072
No	2046/3522 (58.1)	149/235 (63.4)	1897/3287 (57.7)	7.28 (6.16–8.41)	
Yes	1476/3522 (41.9)	86/235 (36.6)	1390/3287 (42.3)	5.83 (4.63–7.02)	
Indoor exposure to biomass[c]					0.0003
No	1861/3524 (52.8)	97/235 (41.3)	1764/3289 (53.6)	5.21 (4.20–6.22)	
Yes	1663/3524 (47.2)	138/235 (58.7)	1525/3289 (46.4)	8.30 (6.97–9.62)	
History of coronary heart disease[c]					0.0469
No	3379/3524 (95.9)	218/235 (92.8)	3161/3289 (96.1)	6.45 (5.62–7.28)	
Yes	145/3524 (95.9)	17/235 (7.2)	128/3289 (3.9)	11.72 (6.49–16.96)	
History of hypertension[c]					0.0393
No	2883/3524 (81.8)	180/235 (76.6)	2703/3289 (82.2)	6.24 (5.36–7.13)	
Yes	641/3524 (18.2)	55/235 (23.4)	586/3289 (17.8)	8.58 (6.41–10.75)	

Definition of abbreviations: CI = confidence interval; PRISm = preserved ratio impaired spirometry; PEF = peak expiratory flow; sSAD = spirometric small airway dysfunction; MMEF = maximal midexpiratory flow; FEF50 = forced expiratory flow at 50% of FVC exhaled; FEF75 = forced expiratory flow at 75% of FVC exhaled. [a]Data are presented as exposed/total participants (%) or mean (SD). [b]P value for the univariate Cox regression in the training cohort. [c]Data missing for chronic bronchitis (n = 2), low PEF (n = 43), sSAD (n = 43), low MMEF (n = 43), low FEF50 (n = 43), low FEF75 (n = 43), emphysema (n = 2), smoking status (n = 5), exposure to dust or chemicals in the workplace (n = 2), indoor exposure to coal (n = 4), indoor exposure to biomass (n = 2), history of coronary heart disease (n = 2), and history of hypertension (n = 2).

Table 1: **Baseline characteristics of 3526 subjects and cumulative incidence of COPD in the training cohort.**

	Crude HR (95% CI)	Model 1[a]		Model 2[b]		Model 3[c]	
		Adjusted HR (95% CI)	P	Adjusted HR (95% CI)	P	Adjusted HR (95% CI)	P
Chronic bronchitis	2.89 (1.40–5.93)	1.80 (0.85–3.85)	0.1274	2.84 (1.38–5.84)	0.0046	2.94 (1.43–6.04)	0.0035
PRISm	1.63 (1.08–2.47)	1.57 (1.03–2.40)	0.0361	0.87 (0.56–1.36)	0.5385	0.76 (0.48–1.19)	0.2270
Low PEF	1.75 (1.32–2.32)	1.66 (1.25–2.21)	0.0005	1.28 (0.94–1.75)	0.1115	1.22 (0.90–1.68)	0.2055
sSAD	2.86 (2.21–3.70)	2.69 (2.07–3.49)	<0.0001	2.74 (2.07–3.61)	<0.0001		
Low MMEF	3.26 (2.50–4.25)	2.93 (2.24–3.83)	<0.0001			2.38 (1.65–3.43)	<0.0001
Low FEF50	2.58 (1.99–3.34)	2.46 (1.89–3.19)	<0.0001			1.42 (0.98–2.06)	0.0622
Low FEF75	1.98 (1.44–2.71)	2.37 (1.73–3.26)	<0.0001			1.23 (0.86–1.76)	0.2598

Definition of abbreviations: CI = confidence interval; PRISm = preserved ratio impaired spirometry; PEF = peak expiratory flow; sSAD = spirometric small airway dysfunction; MMEF = maximal midexpiratory flow; FEF50 = forced expiratory flow at 50% of FVC exhaled; FEF75 = forced expiratory flow at 75% of FVC exhaled. [a]Model 1: included each one of chronic bronchitis, PRISm, low PEF, sSAD, low MMEF, low FEF50, and low FEF75 respectively, adjusted for sex, age, smoking status, exposure to dust or chemicals in the workplace, indoor exposure to coal, and indoor exposure to biomass. [b]Model 2: included chronic bronchitis, PRISm, low PEF, and sSAD. [c]Model 3: included chronic bronchitis, PRISm, low PEF, low MMEF, low FEF50, and low FEF75.

Table 2: **Crude and adjusted hazard ratios of COPD development for potential indicators for pre-COPD.**

18.54/1000 person-years. Subjects who were with chronic bronchitis, PRISm, low PEF, sSAD, low MMEF, low FEF50, and low FEF75 respectively had higher incidence of COPD than those who were without the corresponding indicator.

The relationships between potential indicators/their combinations and COPD development

Each of the seven indicators was associated with an increased risk of developing COPD in the univariate model (Table 2). In the multivariable model for each of them, the association was still significant except for chronic bronchitis after adjustment for COPD-specific risk factors (model 1). The adjusted HRs for PRISm, low PEF, sSAD, low MMEF, low FEF50, and low FEF75 were 1.57 (95% CI 1.03–2.41), 1.65 (1.24–2.21), 2.77 (2.13–3.60), 3.01 (2.30–3.94), 2.52 (1.94–3.28), and 2.34 (1.70–3.21). When only chronic bronchitis, PRISm, low PEF and sSAD were included in model 2, chronic bronchitis and sSAD remained significant. When the place of sSAD was taken by low MMEF, low FEF50 and low FEF75 in model 3, chronic bronchitis and low MMEF remained significant.

As shown in Table 3, combinations of chronic bronchitis and sSAD were categorized into four groups: neither of chronic bronchitis and sSAD, only chronic bronchitis, only sSAD, and both of chronic bronchitis and sSAD. Combinations of chronic bronchitis and low MMEF were also categorized in the same way. Although there was no evidence of significant interaction, both of chronic bronchitis and sSAD and both of chronic bronchitis and low MMEF showed the highest HR of developing COPD compared to the other groups with only one indicator alone.

	Number of cases/participants	Crude HR (95% CI)	P	Number of cases/participants	Adjusted HR (95% CI)[a]	P[a]
Model including the combination of CB and sSAD	231/3481			231/3476		
Neither of CB and sSAD	105/2425	1.00 (ref)		105/2422	1.00 (ref)	
Only CB	4/32	2.94 (1.04–8.33)	0.0427	4/32	1.55 (0.48–4.96)	0.4626
Only sSAD	118/1005	2.86 (2.20–3.72)	<0.0001	118/1003	2.62 (2.00–3.42)	<0.0001
Both of CB and sSAD	4/19	7.74 (2.85–21.05)	<0.0001	4/19	5.79 (2.08–16.08)	0.0008
Model including the interaction item						
CB*sSAD			0.9115			0.6511
Model including the combination of CB and low MMEF	231/3481			231/3476		
Neither of CB and low MMEF	135/2836	1.00 (ref)		135/2832	1.00 (ref)	
Only CB	6/40	3.39 (1.46–7.88)	0.0046	6/40	2.07 (0.83–5.13)	0.1179
Only low MMEF	88/594	3.30 (2.52–4.32)	<0.0001	88/593	2.87 (2.17–3.79)	<0.0001
Both of CB and low MMEF	2/11	7.56 (1.87–30.62)	0.0046	2/11	4.76 (1.15–19.69)	0.0312
Model including the interaction item						
CB*low MMEF			0.6390			0.7996

Definition of abbreviations: CI = confidence interval; CB = chronic bronchitis; sSAD = spirometric small airway dysfunction; MMEF = maximal midexpiratory flow. [a]Adjusted for sex, age, smoking status, exposure to dust or chemicals in the workplace, indoor exposure to coal, indoor exposure to biomass, and history of coronary heart disease and hypertension.

Table 3: **Hazard ratios of COPD development for combinations of chronic bronchitis and sSAD/low MMEF.**

Positive associations were found between any two of PRISm, low PEF, low MMEF, low FEF50, and low FEF75 (eTable 1). When individuals with low MMEF, low FEF50, or chronic bronchitis were excluded for sensitivity analysis, PRISm was not associated with a higher risk of COPD (eTable 2). When individuals with low MMEF or low FEF50 were excluded, low PEF was not associated with a higher risk of COPD (eTable 2).

The prediction model for COPD development

All the variables described in Table 1 were involved in LASSO regression model for screening predictors. SSAD and its three elements were screened in two separate models. Using the minimum standard I-Standard Error method, four predictors were selected for model 4 (age, sex, sSAD, and indoor exposure to biomass) and five for model 5 (age, sex, low MMEF, low FEF50, and indoor exposure to biomass) (eFig. 1 and Table 4). Table 4 presents the results of Cox regression and performance metrics for model 4 and 5. For internal validation, model 5 showed a higher Dxy and C-index but a lower AIC than model 4. The calibration plots for the two models (Fig. 2) demonstrated good concordance between nomogram-predicted probability of developing COPD at 5 years and actual proportion of COPD at 5 years.

eTable 3 presents the baseline characteristics of subjects in the validation cohort. Characteristics were not significantly different between the training cohort and validation cohort except for indoor exposure to coal (1476/3522 (41.91%) versus 132/461 (28.60%)). 30 out of 461 subjects (6.51%) developed COPD in the second survey. For temporal validation, Fig. 3 presents the ROC curves demonstrating the good predictive value of the prognostic index calculated based on the two prediction models (model 4: prognostic index = 0.0382 × age + 0.9977 × male sex + 0.9953 × sSAD + 0.6932 × indoor exposure to biomass; model 5: prognostic index = 0.0394 × age + 0.9206 × male sex + 0.7552 × low MMEF + 0.4541 × low FEF 50 + 0.7169 × indoor exposure to biomass; each variable value except age equaled 1/0 if the corresponding factor was present/absent) and variable values in the validation dataset. However, model 5 showed a higher AUC and Youden's index of the cut-off point (sensitivity + specificity – 1) than model 4. The positive and negative predictive values were also relatively higher in model 5 than in model 4, although these two measures were strongly affected by prevalence of the target disease in a population.[27] The cut-off point could be used to discriminate individuals at a high or low risk of COPD if the prognostic index was calculated. Moreover, Hosmer–Lemeshow test showed good calibration degree for the two models in temporal validation. According to the relatively better internal and temporal performance, model 5 was defined as the final prediction model. The 5-year probability of developing COPD for any individual could be predicted by the nomogram of the final model (Fig. 4).

	Model 4			Model 5		
	Coefficient	SE	P	Coefficient	SE	P
Age	0.0382	0.0069	<0.0001	0.0394	0.0069	<0.0001
Male	0.9977	0.1400	<0.0001	0.9206	0.1419	<0.0001
sSAD	0.9953	0.1325	<0.0001			
Low MMEF				0.7552	0.1812	<0.0001
Low FEF50				0.4541	0.1762	0.0100
Indoor exposure to biomass	0.6932	0.1393	<0.0001	0.7169	0.1387	<0.0001
Internal validation						
Dxy	0.534			0.541		
C	0.7670[a]	0.0171		0.7704[b]	0.0170	
AIC	3184.793			3180.358		
Corrected Dxy	0.527			0.532		
Corrected C	0.7635[c]	0.0228		0.7660[d]	0.0229	
Temporal validation						
AUC	0.715[e]	0.049		0.730[f]	0.050	
X-squared of HL test	1.7459		0.9878	2.8828		0.9415

Definition of abbreviations: sSAD = spirometric small airway dysfunction; MMEF = maximal midexpiratory flow; FEF50 = forced expiratory flow at 50% of FVC exhaled; AIC = akaike information criterion; AUC = area under the receiver operating characteristic curve; HL test = Hosmer–Lemeshow test; SE = standard error. [a]95% Confidence Interval: 0.7335–0.8005. [b]95% Confidence Interval: 0.7371–0.8037. [c]95% Confidence Interval: 0.7188–0.8082. [d]95% Confidence Interval: 0.7210–0.8110. [e]95% Confidence Interval: 0.6190–0.8110. [f]95% Confidence Interval: 0.6320–0.8280.

Table 4: **Cox regression analysis results and performance metrics for the internal and temporal validation of the prediction models for COPD development.**

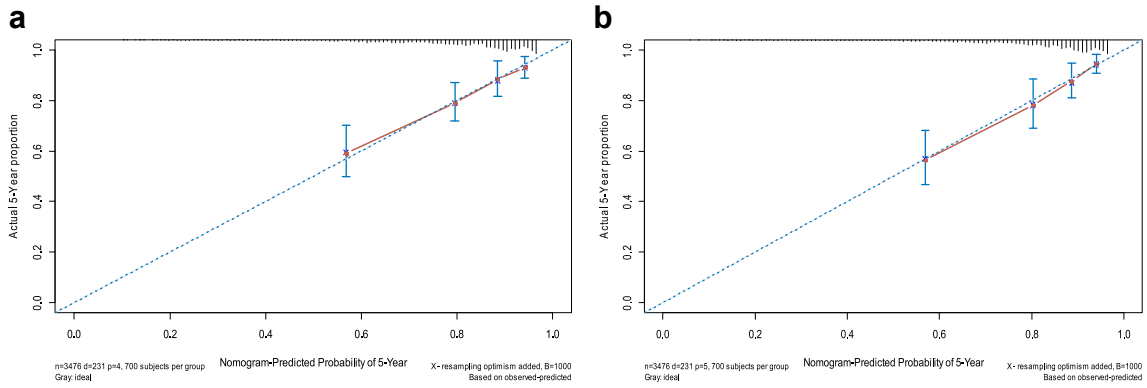

Fig. 2: Calibration plots for the internal validation of the nomograms to predict COPD development (a) for Model 4 and (b) for Model 5). The solid line overlapping with the dashed line demonstrated the concordance between nomogram-predicted probability of developing COPD at 5 years and actual proportion of COPD at 5 years.

Discussion

This study demonstrated that each of PRISm, low PEF, sSAD, low MMEF, low FEF50 and low FEF75 was associated with a higher risk of COPD development and highlighted the effects of chronic bronchitis and sSAD/low MMEF independent of the other indicators. The concurrence of chronic bronchitis and sSAD/low MMEF had the strongest effect and they were the most optimal of those investigated. PRISm was an independent indicator for developing COPD only if low MMEF, low FEF50, and chronic bronchitis were present. Low PEF was an independent indicator only if low MMEF and low FEF50 were present. Our prediction model for COPD development including age, sex, low MMEF, low FEF50, and indoor exposure to biomass had good discrimination and calibration both internally and temporally.

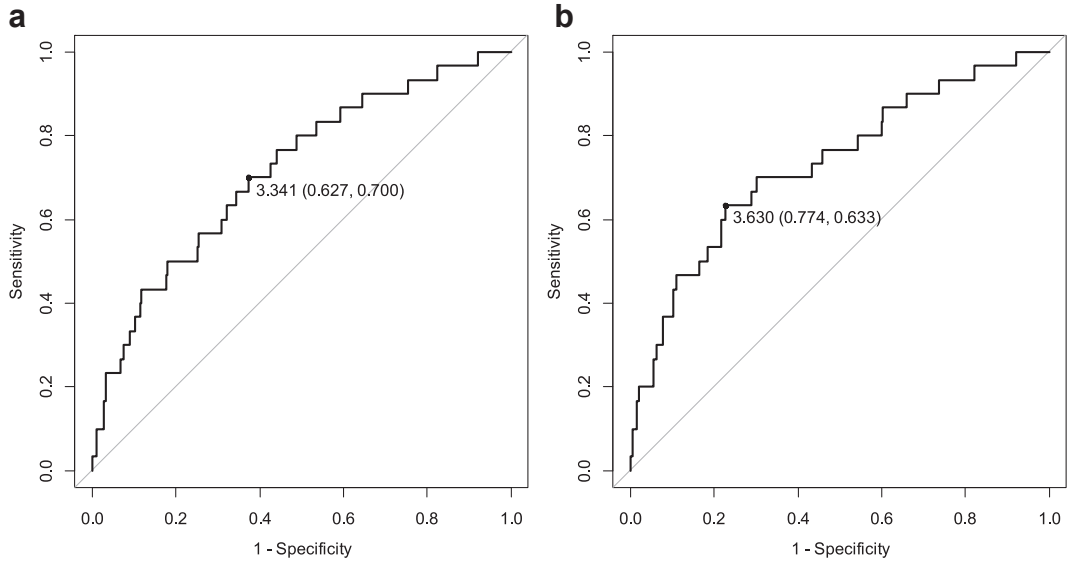

Fig. 3: Receiver operating characteristic curves for the temporal validation of the prediction models for COPD development. a) Model 4: The cut-off value of prognostic index was 3.341 to discriminate individuals at a high or low risk of COPD, at which the sensitivity and the specificity was 0.700 and 0.627 respectively and the positive and negative predictive value was 0.117 and 0.967 respectively. The area under the curve was 0.715; b) Model 5: The cut-off value of prognostic index was 3.630 to discriminate individuals at a high or low risk of COPD, at which the sensitivity and the specificity was 0.633 and 0.774 respectively and the positive and negative predictive value was 0.165 and 0.968 respectively. The area under the curve was 0.730. For example for the usage of the cut-off value in model 5, a 65 year-old (65*0.0394) man (0.9206) with low MMEF (0.7552), but without low FEF50 (0) and indoor exposure to biomass (0) presents a prognostic index of 4.2368 indicating a high risk of COPD.

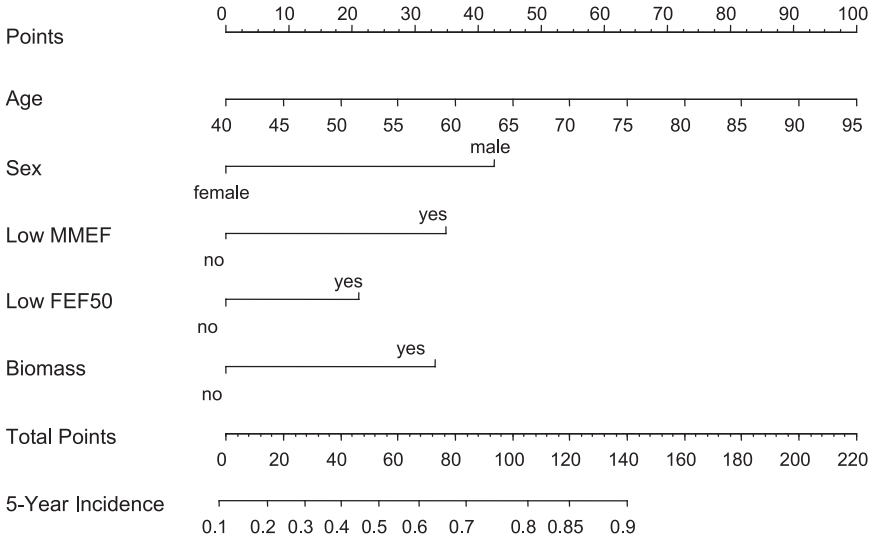

Fig. 4: Nomogram of the final model (model 5) to predict 5-year incidence of COPD. The probability of developing COPD for any individual could be predicted by the straight line drawn from the location of the total points which are obtained by drawing straight lines upward to the points axis for each variable value and summing their points in the nomogram. For example, a 65 year-old (45 points) man (42.5 points) with low MMEF (35 points), but without low FEF50 (0 point) and indoor exposure to biomass (0 point) presents a total of 122.5 points indicating a slightly above 85% probability of developing COPD five years later. Definition of abbreviations: MMEF = maximal mid-expiratory flow; FEF50 = forced expiratory flow at 50% of FVC exhaled.

As small airway could be one of the origins in COPD disease process, small airway dysfunction has been reported as associated with COPD development.[7,8,29,30] However, indicators and outcomes defined in some of these studies were not consistent. Meanwhile, some study populations (e.g., ever smokers, men, and α_1-antitrypsin deficiency patients) were not the general population and their sample size were small, which limited their generalizability. In contrast, the sample in our study represented the general population aged 40 years or older living in China, the results of which could be generalized in Chinese. Our larger sample size also enhanced the evidence of the effect of sSAD and prioritized its three spirometric elements on predicting COPD.

One study in China showed that PEF <80% predicted was a good indicator to screen populations for airflow obstruction,[18] but whether it is effective to predict the future risk of airflow obstruction is uncertain. From the cohort in our study, the adjusted HR for PEF <80% predicted suggested its value in prediction for COPD besides screening. This implicated that even if subjects with low PEF had normal FEV_1/FVC temporarily, they were still at higher probability of developing COPD in years. Several studies have reported the higher percentage of individuals who would develop COPD in years among those with PRISm compared to healthy controls.[17,31–33] In our study the adjusted HR for PRISm was relatively lower than the ones reported previously,[34,35] which was probably caused by the environmental factors adjusted for in our models.

In line with one study in the US,[36] the association of chronic bronchitis with COPD development was significant in univariate analysis but not significant when other factors were taken into account. However, there were others reporting significant associations in multivariable analysis.[4,37,38] The inconsistent results may be attributed to different covariates which could be confounding factors between chronic bronchitis and COPD.

When chronic bronchitis, PRISm, low PEF and sSAD were included in the model, only chronic bronchitis and sSAD remained significant indicating that these two might progress to COPD in different processes. This was consistent with a previous study putting forward two trajectories of progression to COPD.[39] When sSAD was replaced by its three elements, low MMEF implied the most valuable early-warning sign for COPD compared to the other two. Although there have been different methods for assessing small airway dysfunction, most of them are difficult to perform, costly, not widely available, with exposure to radiation or in need of further studies.[20] By comparison, spirometry is a noninvasive method and relatively easy to perform. Given that there is no universally accepted indicator for small airway dysfunction, our study found the best one in predicting COPD among several spiromtric parameters. In addition, the HR for the coexistence of chronic bronchitis and sSAD/low MMEF was higher than the HR for the existence of each of them alone, suggesting the prominence of their concurrence for the first time and that eliminating either of them was important to reduce the risk of the development of COPD.

As COPD can be described from symptom, structure, function, and causation, we intended to predict the development of COPD from different perspectives as well. Although previous COPD prediction models involved common risk factors for COPD,[11] they did not involve potential indicators for pre-COPD. Our study took advantage of screened potential clinical indicators for pre-COPD that could be easily obtained to predict COPD from function, causation, and demography more comprehensively. Chronic bronchitis, smoking, and indoor exposure to coal were not included probably because they were confounded by the independent predictors in the model. As pathological changes can persist on smoking cessation,[2] the improvement of spirometric parameters such as MMEF and FEF50 have to be achieved when eliminating smoking or other environmental risk factors in order to prevent COPD effectively. The good performance in both the internal and temporal validation indicated its consistent value in prediction.

Within our study, PRISm and low PEF were substantially associated with the three elements of sSAD among persons without COPD. The positive association between PRISm and small airway dysfunction was reported in two previous studies consistently.[40,41] Similarly, PEF was positively associated with most parameters of small airway function in children.[42] In addition to the associations we observed between indicators at baseline, sensitivity analysis further revealed that PRISm was an independent indicator for developing COPD only if low MMEF, low FEF50, and chronic bronchitis were present, and low PEF was an independent indicator only if low MMEF and low FEF50 were present, which indicated the pathophysiological mechanism underlying them. This was consistent with their significant HRs in their respective model but non-significant HRs when all the indicators were included at once. This finding also suggested that the association of PRISm or low PEF with increased risk of COPD development was not as generalizable as that reported in previous studies but conditional to some extent.

A major strength of this study is that, it involved more than one indicator proposed for pre-COPD in one cohort making comparisons of their effects more evident. To our knowledge, this study is the first one to provide prospective evidence for it based on a cohort of subjects across six regions of China making our results generalizable to the Chinese population. Second, we also examined their independent and additive effects when they were involved simultaneously or combined for the first time. Third, considering potential indicators for pre-COPD from specific refined clinical features made our prediction model more comprehensive and concise. Fourth, we examined the associations among potential indicators contemporarily and longitudinally providing a clue to understand the mechanisms of COPD development. Interventions and treatments could be implemented targeting the mechanisms underlying the indicators.

Our study also has limitations. First, only two time points of spirometry and a short follow-up period were available, so variability in the measurements might contaminate the definition of incident COPD. Second, the spirometric indicators were selected according to previous studies on predicting or screening COPD. Indicators of sSAD were those which had already been frequently used to assess the function of small airways and measured in our surveys. Those which were seldom reported as predictors of COPD or used as indicators of sSAD, or having few established methods for calculating predicted values, were not included in our study. For example, a novel spirometric measure the Peak Index (the number of peaks adjusted for lung size) which had been found significantly associated with small airway disease was not involved.[43] Third, since the participants in our study are all Chinese residents aged 40 years or older, which is a high risk population for COPD in terms of age, the applicability to populations in other countries or younger populations is limited. The sample size of the temporal validation dataset was relatively small compared to the training dataset, and the results could be validated within larger samples in future studies.

In conclusion, each of PRISm, low PEF, sSAD, low MMEF, low FEF50 and low FEF75 was associated with a higher risk of COPD development and the effects of chronic bronchitis and sSAD/low MMEF were independent of the other indicators. The concurrence of chronic bronchitis and sSAD/low MMEF had the strongest effect and they were the most optimal of those investigated. Physicians should pay more attention to patients with chronic bronchitis and sSAD/low MMEF concurrently. The effect of PRISm on increasing risk of COPD was dependent on low MMEF, low FEF50, and chronic bronchitis. The effect of low PEF on increasing risk of COPD was dependent on low MMEF and low FEF50. Our prediction model had good performance and could serve as a multi-dimension tool with the prognostic index and nomogram to predict COPD more comprehensively. Further studies are needed to illustrate the particular risks of COPD in populations with or without exposures to specific COPD risk factors and develop specific prediction models to understand the natural history of COPD in various contexts and facilitate interventions and treatments accordingly.

Contributors

LF and YC developed the research aim and study design. JF analyzed the data, drafted the paper and was involved in the study design. LF, JF, SC, YZ, XJ, and NW contributed to data collection, preparation, analysis and interpretation. All authors critically reviewed the manuscript and approved the final version before submission.

Data sharing statement

The data that support the findings of this study are available from National Center for Chronic and Non-communicable Disease Control and

Prevention, Chinese Center for Disease Control and Prevention (China CDC) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of China CDC. Data access requests should be addressed to the National Center for Chronic and Non-communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention (China CDC), Xicheng District, 100050, Beijing, China.

Declaration of interests

We declare no competing interests.

Acknowledgements

We thank all research staff from local Centers for Disease Control and Prevention and local hospitals for collection of data. This study was supported by grants from the Ministry of Finance of the People's Republic of China (Promotion and Administration of COPD Surveillance System), the Ministry of Science and Technology of the People's Republic of China (National Key R&D Programme of China, 2016YFC1303905), and the National Natural Science Foundation of China (grant no 82090014). The funding sources of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report or the decision to submit this report for publication.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanwpc.2023.100984.

References

1 World Health Organization. Global health estimates: life expectancy and leading causes of death and disability. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates. Accessed July 7, 2022.
2 Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (2023 report). http://goldcopd.org. Accessed November 22, 2022.
3 Han MK, Agusti A, Celli BR, et al. From GOLD 0 to pre-COPD. *Am J Respir Crit Care Med*. 2021;203(4):414–423.
4 Lindberg A, Jonsson AC, Rönmark E, Lundgren R, Larsson LG, Lundbäck B. Ten-year cumulative incidence of COPD and risk factors for incident disease in a symptomatic cohort. *Chest*. 2005;127(5):1544–1552.
5 Allinson JP, Hardy R, Donaldson GC, Shaheen SO, Kuh D, Wedzicha JA. The presence of chronic mucus hypersecretion across adult life in relation to chronic obstructive pulmonary disease development. *Am J Respir Crit Care Med*. 2016;193(6):662–672.
6 Park HJ, Byun MK, Rhee CK, Kim K, Kim HJ, Yoo KH. Significant predictors of medically diagnosed chronic obstructive pulmonary disease in patients with preserved ratio impaired spirometry: a 3-year cohort study. *Respir Res*. 2018;19(1):185.
7 Yee N, Markovic D, Buhr RG, et al. Significance of FEV(3)/FEV(6) in recognition of early airway disease in smokers at risk of development of COPD: analysis of the SPIROMICS cohort. *Chest*. 2022;161(4):949–959.
8 Harvey BG, Strulovici-Barel Y, Kaner RJ, et al. Risk of COPD with obstruction in active smokers with normal spirometry and reduced diffusion capacity. *Eur Respir J*. 2015;46(6):1589–1597.
9 Oelsner EC, Smith BM, Hoffman EA, et al. Associations between emphysema-like lung on CT and incident airflow limitation: a general population-based cohort study. *Thorax*. 2018;73(5):486–488.
10 Zeng S, Tham A, Bos B, Jin J, Giang B, Arjomandi M. Lung volume indices predict morbidity in smokers with preserved spirometry. *Thorax*. 2019;74(2):114–124.
11 Matheson MC, Bowatte G, Perret JL, et al. Prediction models for the development of COPD: a systematic review. *Int J Chron Obstruct Pulmon Dis*. 2018;13:1927–1935.
12 Perret JL, Vicendese D, Simons K, et al. Ten-year prediction model for post-bronchodilator airflow obstruction and early detection of COPD: development and validation in two middle-aged population-based cohorts. *BMJ Open Respir Res*. 2021;8(1):e001138.
13 Fang L, Gao P, Bao H, et al. Chronic obstructive pulmonary disease in China: a nationwide prevalence study. *Lancet Respir Med*. 2018;6(6):421–430.
14 Asthma Committee, Chinese Thoracic Society. Guidelines for the diagnosis and treatment of cough (2009). *Chin J Tuberc Respir Dis*. 2009;32(6):407–413.
15 Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J*. 2005;26(5):948–968.
16 Stanojevic S, Kaminsky DA, Miller MR, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J*. 2022;60(1):2101499.
17 Higbee DH, Granell R, Davey Smith G, Dodd JW. Prevalence, risk factors, and clinical implications of preserved ratio impaired spirometry: a UK Biobank cohort analysis. *Lancet Respir Med*. 2022;10(2):149–157.
18 Tian J, Zhou Y, Cui J, et al. Peak expiratory flow as a screening tool to detect airflow obstruction in a primary health care setting. *Int J Tuberc Lung Dis*. 2012;16(5):674–680.
19 Xiao D, Chen Z, Wu S, et al. Prevalence and risk factors of small airway dysfunction, and association with smoking, in China: findings from a national cross-sectional study. *Lancet Respir Med*. 2020;8(11):1081–1093.
20 Li Y, Li XY, Yuan LR, Wang HL, Pang M. Evaluation of small airway function and its application in patients with chronic obstructive pulmonary disease (review). *Exp Ther Med*. 2021;22(6):1386.
21 Jian W, Gao Y, Hao C, et al. Reference values for spirometry in Chinese aged 4-80 years. *J Thorac Dis*. 2017;9(11):4538–4549.
22 Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. *Eur Respir J*. 1993;6 Suppl 16:5–40.
23 Tang XR, Li YQ, Liang SB, et al. Development and validation of a gene expression-based signature to predict distant metastasis in locoregionally advanced nasopharyngeal carcinoma: a retrospective, multicentre, cohort study. *Lancet Oncol*. 2018;19(3):382–393.
24 Zhou ZR, Wang WW, Li Y, et al. In-depth mining of clinical data: the construction of clinical prediction model with R. *Ann Transl Med*. 2019;7(23):796.
25 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361–387.
26 Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Stat Med*. 1990;9(11):1303–1325.
27 Hajian-Tilaki K. The choice of methods in determining the optimal cut-off value for quantitative diagnostic test evaluation. *Stat Methods Med Res*. 2018;27(8):2374–2383.
28 Lemeshow S, Hosmer DW, Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol*. 1982;115(1):92–106.
29 Stockley JA, Ismail AM, Hughes SM, Edgar R, Stockley RA, Sapey E. Maximal mid-expiratory flow detects early lung disease in α(1)-antitrypsin deficiency. *Eur Respir J*. 2017;49(3):1602055.
30 Olofson J, Bake B, Bergman B, Vanfleteren L, Svärdsudd K. Prediction of COPD by the single-breath nitrogen test and various respiratory symptoms. *ERJ Open Res*. 2021;7(3):00383-2021.
31 Wan ES, Fortis S, Regan EA, et al. Longitudinal phenotypes and mortality in preserved ratio impaired spirometry in the COPDGene study. *Am J Respir Crit Care Med*. 2018;198(11):1397–1405.
32 Lange P, Celli B, Agustí A, et al. Lung-function trajectories leading to chronic obstructive pulmonary disease. *N Engl J Med*. 2015;373(2):111–122.
33 Wijnant SRA, De Roos E, Kavousi M, et al. Trajectory and mortality of preserved ratio impaired spirometry: the Rotterdam Study. *Eur Respir J*. 2020;55(1):1901217.
34 Washio Y, Sakata S, Fukuyama S, et al. Risks of mortality and airflow limitation in Japanese individuals with preserved ratio impaired spirometry. *Am J Respir Crit Care Med*. 2022;206(5):563–572.
35 Kanetake R, Takamatsu K, Park K, Yokoyama A. Prevalence and risk factors for COPD in subjects with preserved ratio impaired spirometry. *BMJ Open Respir Res*. 2022;9(1):e001298.

36 Guerra S, Sherrill DL, Venker C, Ceccato CM, Halonen M, Martinez FD. Chronic bronchitis before age 50 years predicts incident airflow limitation and mortality risk. *Thorax*. 2009;64(10):894–900.

37 Probst-Hensch NM, Curjuric I, Pierre-Olivier B, et al. Longitudinal change of prebronchodilator spirometric obstruction and health outcomes: results from the SAPALDIA cohort. *Thorax*. 2010;65(2):150–156.

38 de Marco R, Accordini S, Cerveri I, et al. Incidence of chronic obstructive pulmonary disease in a cohort of young adults according to the presence of chronic cough and phlegm. *Am J Respir Crit Care Med*. 2007;175(1):32–39.

39 Young AL, Bragman FJS, Rangelov B, et al. Disease progression modeling in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2020;201(3):294–302.

40 Zhao N, Wu F, Peng J, et al. Preserved ratio impaired spirometry is associated with small airway dysfunction and reduced total lung capacity. *Respir Res*. 2022;23(1):298.

41 Ding Q, Mi BB, Wei X, et al. Small airway dysfunction in chronic bronchitis with preserved pulmonary function. *Can Respir J*. 2022;2022:4201786.

42 Wu XJ, Huang Y, Wang Y, Gong CH. [Association of FEV1 and PEF with small airway function in asthmatic children: cross-sectional analysis of 619 cases]. *Nan Fang Yi Ke Da Xue Xue Bao*. 2011;31(11):1900–1902.

43 Bhatt SP, Bodduluri S, Raghav V, et al. The peak index: spirometry metric for airflow obstruction severity and heterogeneity. *Ann Am Thorac Soc*. 2019;16(8):982–989.