

Association between low peak expiratory flow and newly reported physician-diagnosed chronic lung disease among middle-aged and older Chinese adults: a prospective cohort study

Hao Haohao#

Fourth Affiliated Hospital of Harbin Medical University

Chen Chen

Daqing City People's Hospital

Yunhao Jiang

1019182985@qq.com

Fourth Affiliated Hospital of Harbin Medical University

Research Article

Keywords: peak expiratory flow, chronic lung disease, CHARLS, prospective cohort

Posted Date: September 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10707526/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Peak expiratory flow (PEF) is an inexpensive measure of expiratory performance, but whether it identifies respiratory vulnerability before chronic lung disease is diagnosed remains uncertain. We examined whether PEF below values expected for age, sex, and height was associated with a newly reported physician diagnosis of chronic lung disease.

Methods

This prospective cohort study included 10,031 participants aged 45 years or older from the China Health and Retirement Longitudinal Study (CHARLS). All had valid 2011 baseline PEF measurements, completed at least one follow-up disease assessment, and reported neither physician-diagnosed chronic lung disease nor asthma at baseline. Low PEF was defined as a sex-specific standardized residual below -1.282 , derived from reference equations incorporating age, age squared, and height. The outcome was the first newly reported physician diagnosis of chronic lung disease from 2013 to 2020. We estimated associations using complementary log-log discrete-time survival models and calculated standardized 9-year risks. Additional Cox, spline, alternative-threshold, weighting, gap-corrected, and landmark analyses assessed robustness.

Results

Of 10,031 participants, 1,068 (10.6%) had low PEF. During a mean follow-up of 7.77 years, 1,243 participants first reported physician-diagnosed chronic lung disease. In the fully adjusted model, low PEF was associated with a higher hazard of the outcome (hazard ratio [HR] 1.406, 95% confidence interval [CI] 1.193–1.657). The standardized 9-year risk was 19.0% (95% CI 16.8–21.9%) under low PEF and 14.0% (95% CI 13.2–14.8%) under non-low PEF. The adjusted risk difference was 5.1 percentage points (95% CI 2.6–8.1). Estimates were similar in the gap-corrected analysis (HR 1.405, 95% CI 1.192–1.655) and after inverse probability of censoring weighting (HR 1.445, 95% CI 1.209–1.726). The estimate was attenuated in the 2015 landmark analysis (HR 1.255, 95% CI 0.983–1.601).

Conclusions

PEF below values expected for age, sex, and height was associated with a higher risk of subsequently reported physician-diagnosed chronic lung disease. Attenuation of this association after a longer diagnostic lag suggests that low PEF captures a critical window of underlying, unrecognized respiratory impairment years before formal diagnosis. Therefore, PEF may serve as a valuable early warning tool in

settings where spirometry is underutilized. Because the outcome relied on reported diagnostic status, these findings do not establish spirometry-confirmed chronic obstructive pulmonary disease.

1.Introduction

Chronic respiratory diseases are major causes of disability and death worldwide, with chronic obstructive pulmonary disease (COPD) accounting for a substantial share of this burden [1]. The China Pulmonary Health study identified post-bronchodilator spirometric COPD in 8.6% of adults aged 20 years or older [2]. However, respiratory disease may remain unrecognized until symptoms or functional impairment prompt clinical evaluation. Measures of respiratory performance may therefore identify vulnerability before a diagnosis is reported. Nevertheless, COPD diagnosis requires evidence of persistent post-bronchodilator airflow obstruction [3].

Peak expiratory flow (PEF) is portable, inexpensive, and feasible in community and primary care settings. It reflects airway caliber and lung volume but is also influenced by expiratory muscle strength, effort, age, sex, and body size. Chinese reference data show substantial variation in PEF according to age, sex, and height [5]. Fixed cutoffs may therefore misclassify individuals whose expected physiological values differ. A standardized residual (SR) expresses the difference between observed and predicted PEF relative to residual variation [6].

Low PEF is included in COPD case-finding strategies, such as CAPTURE, to identify adults who may benefit from diagnostic testing [7]. Longitudinal studies have associated low PEF with health-care use, disability, frailty, and mortality [8–10]. Analyses of the China Health and Retirement Longitudinal Study (CHARLS) have linked sarcopenia, circadian syndrome, and joint PEF-muscle-mass trajectories to subsequent chronic lung disease [11–13]. However, whether an individualized low-PEF phenotype precedes a new physician diagnosis among adults without reported chronic respiratory disease remains unclear.

We therefore examined the association between low baseline PEF and the first newly reported physician diagnosis of chronic lung disease over nine years in CHARLS. We estimated relative and absolute risks, characterized the dose-response association, and tested robustness across alternative PEF definitions and approaches to missing follow-up data. We also excluded diagnoses reported early in follow-up to assess the potential influence of unrecognized disease at baseline.

2.Methods

2.1Study design

This prospective cohort study used the 2011 CHARLS baseline survey and follow-up data from 2013, 2015, 2018, and 2020. CHARLS enrolled community-dwelling Chinese adults aged 45 years or older and their spouses using multistage probability sampling. The cohort design and field procedures have been

described previously [4]. We reported this study according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [14].

2.2 Study population

The 2011 CHARLS analytic release included a total of 17,705 participants. We excluded 433 participants who were younger than 45 years or had unknown sex and 3,598 without linked baseline biomarker records. We then excluded 1,473 who did not meet PEF and height quality criteria. Of these, 694 had fewer than two valid PEF readings, 636 lacked interviewer-confirmed full effort, and 143 had missing or implausible height. Further exclusions were 1,508 participants with physician-diagnosed chronic lung disease or asthma at baseline, 116 with malignant tumor, and 416 without any valid follow-up outcome assessment. This process yielded 10,161 participants eligible for the PEF and outcome analyses. After excluding 130 participants with missing data for at least one prespecified covariate, 10,031 remained in the complete-case analysis (Fig. 1).

2.3 Peak expiratory flow assessment

PEF was measured three times at baseline with a handheld peak flow meter and recorded in liters per minute. Values coded as 993 or 999 were treated as missing, and measurements outside 30–800 L/min were considered implausible. We required at least two valid measurements and interviewer-confirmed full effort. The highest valid reading was defined as maximum PEF (PEFmax), consistent with previous CHARLS research [5]. Test position was retained for quality control but was not included in the prespecified adjustment set.

We fitted sex-specific reference equations among participants with valid PEF who had never smoked and reported no baseline chronic lung disease, asthma, malignant tumor, heart disease, or stroke. The reference samples comprised 1,170 men and 4,476 women. Within each sex, PEFmax was regressed on age, age squared, and measured height. Model coefficients, residual standard deviations, and coefficients of determination are reported in Supplementary Table S7. The PEF SR was calculated as follows:

$$\text{PEF SR} = (\text{observed PEF} - \text{predicted PEF}) / \text{residual standard deviation}.$$

Low PEF, the primary exposure, was defined as a PEF SR below -1.282 , the theoretical 10th percentile of a standard normal distribution. This threshold was not the empirical 10th percentile of the analytic cohort. Participants with a PEF SR of -1.282 or higher formed the reference group. We also analyzed PEF SR continuously and expressed associations per 1-SD lower PEF. Alternative definitions were a PEF SR below the 5th percentile and PEF below 80% predicted. We also applied CAPTURE thresholds of 350 L/min or lower in men and 250 L/min or lower in women [7].

2.4 Outcome assessment

The outcome was the first newly reported physician diagnosis of chronic lung disease after baseline. At each follow-up, participants were asked whether a doctor had diagnosed chronic lung disease; examples included chronic bronchitis and emphysema. Questionnaire wording, preloaded fields, and routing differed across survey waves. We therefore harmonized current-wave and update variables separately for 2013, 2015, 2018, and 2020 using the corresponding codebooks. The first follow-up wave with an affirmative response was defined as the event wave.

The primary person-period dataset comprised four scheduled intervals: 2011–2013, 2013–2015, 2015–2018, and 2018–2020. Participants contributed observations until their first positive report or last valid disease assessment. If a scheduled assessment was missing but a later assessment was available, the later assessment was retained in the primary dataset. A sensitivity analysis reconstructed each interval from the last observed negative assessment to the next observed assessment. This model included the logarithm of the actual interval duration as an offset. Pulmonary function confirmation was unavailable, and the outcome was self-reported. We therefore interpreted it as a newly reported physician diagnosis of chronic lung disease, not spirometry-confirmed incident COPD.

2.5Covariates

Nine baseline covariates were prespecified: age, sex, education, marital status, smoking, alcohol consumption, body mass index (BMI), hypertension, and diabetes. When included in a model, age and BMI were specified as continuous covariates. Education was classified as elementary school or below versus middle school or above. Marital status was classified as married and living with a spouse, married but not living with a spouse, or single, divorced, or widowed. Ever smoking was defined as having smoked at any time. Regular or occasional alcohol consumption during the previous year was classified as drinking.

BMI was calculated as measured weight in kilograms divided by height in meters squared. Weight outside 30–200 kg, height outside 120–210 cm, and BMI outside 10–60 kg/m² were considered implausible. Hypertension was defined by a self-reported physician diagnosis, antihypertensive treatment, mean systolic blood pressure of 140 mmHg or higher, or mean diastolic blood pressure of 90 mmHg or higher. Mean blood pressure was calculated when at least two measurements were available. Diabetes was defined by a self-reported physician diagnosis, glucose-lowering treatment, or qualifying laboratory values. Laboratory criteria were glycated hemoglobin of 6.5% or higher, fasting plasma glucose of 126 mg/dL or higher, or nonfasting glucose of 200 mg/dL or higher.

2.6Statistical analysis

Continuous variables were summarized as mean (SD), and categorical variables as number (percentage). We assessed between-group balance using standardized mean differences (SMDs), with an absolute SMD of 0.10 or greater considered potentially meaningful. Covariates were prespecified rather than selected from baseline group differences.

The primary analysis used generalized linear models with a binomial distribution and complementary log-log link. Follow-up interval entered each model as a categorical time variable. Exponentiated coefficients are reported as discrete-time hazard ratios (HRs) with 95% confidence intervals (CIs). Participant-clustered robust covariance estimates accounted for repeated observations within individuals. Model 1 included PEF group and categorical follow-up interval terms. Model 2 additionally included age, sex, marital status, and education. Model 3 further included smoking, alcohol consumption, BMI, hypertension, and diabetes and served as the fully adjusted model.

We estimated standardized cumulative risks directly from the fully adjusted Model 3. Low-PEF status was assigned first as 0 and then as 1 for every participant while observed covariates were retained. Predicted interval-specific hazards were converted to cumulative risks and averaged across the cohort. The 9-year risk difference was the standardized risk under low PEF minus the risk under non-low PEF. Model-based 95% CIs were obtained from 500 simulations using the participant-clustered covariance matrix.

We examined the association across the PEF distribution using a restricted cubic spline with four knots at the 5th, 35th, 65th, and 95th percentiles. A PEF SR value of 0 served as the model reference. The spline model included follow-up interval and all Model 3 covariates. Wald tests assessed both the overall association and its nonlinear component. We interpreted a non-significant nonlinearity test as limited evidence of departure from linearity, not as proof of a linear association.

We performed several prespecified sensitivity analyses to evaluate the robustness of the primary findings. First, the gap-corrected model defined intervals from the last observed negative assessment to the next disease assessment and included log interval duration as an offset. Second, a 2013 landmark analysis included participants who reported no disease in 2013 and began follow-up after that wave. A stricter 2015 landmark analysis required negative assessments in both 2013 and 2015 and began follow-up after 2015. Third, pooled logistic models estimated inverse probability of censoring weights (IPCW) for observation at each wave using the nine baseline covariates. For this analysis, follow-up was censored at the first missing disease assessment. Stabilized weights were truncated at the 1st and 99th percentiles, and the primary model was refitted with these weights [15].

Cox proportional hazards models approximated event time as the midpoint between the last disease-free interview and the first positive interview. Participants without an event were censored at their last valid disease assessment; available death timing was used when present. We assessed the proportional hazards assumption in each Cox model using Schoenfeld residuals. We repeated the discrete-time model for each alternative PEF definition. Stabilized inverse probability of treatment weights (IPTW) were derived from a logistic propensity-score model containing all nine covariates. We truncated weights at the 1st and 99th percentiles and applied them in a robust Cox model [15]. IPTW was used to balance measured baseline characteristics between PEF groups, not to adjust for loss to follow-up. Finally, we fitted Cox models within strata of sex, smoking, alcohol consumption, and obesity (BMI below 28 vs. 28 kg/m² or higher). Multiplicative interaction terms tested for heterogeneity across the prespecified strata.

We did not impute missing PEF measurements or outcome data. The primary analysis included only participants with complete data for all nine covariates. Missing covariate data excluded 130 of 10,161 eligible participants (1.3%). All tests were two-sided, and $P < 0.05$ was considered statistically significant. Non-significant interaction P values were interpreted as insufficient evidence of heterogeneity, not as evidence of equivalence. All analyses were performed using R statistical software, version 4.5.2.

3.Results

3.1 Baseline characteristics

The final cohort comprised 10,031 participants, including 1,068 (10.6%) who met the primary definition of low PEF (Fig. 1). Mean age was 58.63 years (SD 9.18), and 5,297 participants (52.8%) were women. Compared with participants without low PEF, those with low PEF had lower mean PEFmax (140.64 vs. 328.10 L/min) and PEF SR (-1.76 vs. 0.17). Age showed the largest imbalance among adjustment variables (SMD 0.143); all other prespecified covariates had SMDs below 0.10 (Table 1).

Table 1
Baseline characteristics of participants according to PEF status

	Overall	PEF status		SMD
		Non-low	Low	
n	10,031	8,963	1,068	
Age, years	58.63 (9.18)	58.76 (9.26)	57.51 (8.34)	0.143
Sex, n (%)				
Male	4,734 (47.2)	4,214 (47.0)	520 (48.7)	0.034
Female	5,297 (52.8)	4,749 (53.0)	548 (51.3)	
Marital status, n (%)				
Married and living with spouse	8,417 (83.9)	7,518 (83.9)	899 (84.2)	0.035
Married but not living with spouse	446 (4.4)	405 (4.5)	41 (3.8)	
Single, divorced, or widowed	1,168 (11.6)	1,040 (11.6)	128 (12.0)	
Education, n (%)				
Elementary school or below	6,737 (67.2)	5,987 (66.8)	750 (70.2)	0.074
Middle school or above	3,294 (32.8)	2,976 (33.2)	318 (29.8)	
Smoking status, n (%)				
Yes	3,868 (38.6)	3,428 (38.2)	440 (41.2)	0.060
No	6,163 (61.4)	5,535 (61.8)	628 (58.8)	
Drinking status, n (%)				
Yes	3,388 (33.8)	3,000 (33.5)	388 (36.3)	0.060
No	6,643 (66.2)	5,963 (66.5)	680 (63.7)	
BMI, kg/m²	23.63 (3.81)	23.66 (3.79)	23.30 (4.02)	0.093
Hypertension, n (%)				
Yes	3,983 (39.7)	3,537 (39.5)	446 (41.8)	0.047
Continuous variables are presented as mean (SD), and categorical variables as n (%). For clarity, SMDs are presented once for each multilevel categorical characteristic. Low PEF was defined as an age-, sex-, and height-standardized residual below - 1.282.				
<i>BMI, body mass index; PEF, peak expiratory flow; SD, standard deviation; SMD, standardized mean difference.</i>				

	Overall	PEF status		SMD
		Non-low	Low	
No	6,048 (60.3)	5,426 (60.5)	622 (58.2)	
Diabetes, n (%)				
Yes	1,226 (12.2)	1,087 (12.1)	139 (13.0)	0.027
No	8,805 (87.8)	7,876 (87.9)	929 (87.0)	
PEFmax, L/min	308.14 (123.71)	328.10 (113.33)	140.64 (60.02)	2.060
PEF standardized residual	-0.03 (1.01)	0.17 (0.84)	-1.76 (0.44)	2.873
Continuous variables are presented as mean (SD), and categorical variables as n (%). For clarity, SMDs are presented once for each multilevel categorical characteristic. Low PEF was defined as an age-, sex-, and height-standardized residual below - 1.282.				
<i>BMI, body mass index; PEF, peak expiratory flow; SD, standard deviation; SMD, standardized mean difference.</i>				

3.2 Association between low PEF and newly reported chronic lung disease

During a mean follow-up of 7.77 years (SD 2.26; median 9.00 years), 1,243 participants first reported physician-diagnosed chronic lung disease. Across follow-up, the cohort contributed 77,925 person-years of observation. Events occurred in 170 of 1,068 participants with low PEF (15.9%) and 1,073 of 8,963 without low PEF (12.0%). Approximate incidence rates were 21.30 and 15.34 per 1,000 person-years, respectively. First positive reports occurred in 2013 for 164 participants, 2015 for 291, 2018 for 464, and 2020 for 324.

Low PEF was associated with a higher hazard of newly reported chronic lung disease in all three discrete-time models (Table 2; Fig. 3). The HR was 1.400 (95% CI 1.189–1.648) in Model 1 and 1.414 (95% CI 1.201–1.666) in Model 2. In fully adjusted Model 3, the HR was 1.406 (95% CI 1.193–1.657; $P < 0.001$).

Table 2
Association between low PEF and newly reported physician-diagnosed chronic lung disease

Model	HR (95% CI)	P value
Model 1	1.400 (1.189–1.648)	< 0.001
Model 2	1.414 (1.201–1.666)	< 0.001
Model 3	1.406 (1.193–1.657)	< 0.001
HRs were estimated with complementary log-log discrete-time survival models using participant-clustered robust standard errors. Model 1 included low PEF status and categorical follow-up interval. Model 2 additionally adjusted for age, sex, marital status, and education. Model 3 further adjusted for smoking, alcohol consumption, BMI, hypertension, and diabetes. CI, confidence interval; HR, hazard ratio; PEF, peak expiratory flow.		

Under low PEF, the standardized 9-year cumulative risk was 19.0% (95% CI 16.8–21.9%). Under non-low PEF, it was 14.0% (95% CI 13.2–14.8%) (Fig. 2; Table 3). This corresponded to an adjusted risk difference of 5.1 percentage points (95% CI 2.6–8.1).

Table 3
Standardized 9-year cumulative risk according to PEF status

Exposure	Standardized 9-year risk, % (95% CI)	Risk difference, percentage points (95% CI)
Non-low PEF	14.0 (13.2–14.8)	Reference
Low PEF	19.0 (16.8–21.9)	5.1 (2.6–8.1)
Risks were standardized to the covariate distribution of the complete analytic cohort using Model 3. Model-based 95% CIs were obtained from 500 simulations using the participant-clustered covariance matrix. CI, confidence interval; PEF, peak expiratory flow.		

3.3Dose-response association

When modeled continuously, each 1-SD lower PEF SR was associated with a 20.7% higher hazard. The corresponding HR was 1.207 (95% CI 1.139–1.280; $P < 0.001$; Supplementary Table S2). The restricted cubic spline supported an overall association across the PEF SR distribution ($P < 0.001$). Evidence of departure from linearity was limited (P for nonlinearity = 0.107; Fig. 5).

3.4Subgroup and sensitivity analyses

The midpoint-time Cox analysis yielded an estimate similar to the primary discrete-time model. The fully adjusted HR was 1.404 (95% CI 1.193–1.652; $P < 0.001$; Supplementary Table S1). The proportional hazards assumption was not rejected for low PEF ($P = 0.099$) or globally ($P = 0.378$; Supplementary Table S5).

Associations remained consistent across all prespecified alternative exposure definitions evaluated. The fully adjusted HR was 1.462 (95% CI 1.169–1.828) for PEF SR below the 5th percentile. For PEF below

80% predicted, the HR was 1.350 (95% CI 1.199–1.520). The HR for the CAPTURE threshold was 1.386 (95% CI 1.229–1.562). All alternative-definition associations had P values below 0.001 in fully adjusted models (Supplementary Table S2).

After IPTW, the effective sample size was 9,978, and the maximum absolute SMD decreased from 0.143 to 0.033. The weighted HR was 1.389 (95% CI 1.176–1.639; $P < 0.001$; Supplementary Table S4).

Overall, 1,492 participants had a later observed disease assessment after an earlier missing assessment, and 155 first positive reports followed such a gap. The gap-corrected analysis retained 10,031 participants and 1,243 events, yielding an HR of 1.405 (95% CI 1.192–1.655; $P < 0.001$). The 2013 landmark analysis included 8,762 participants with a negative 2013 assessment and 1,012 subsequent events. Its HR was 1.344 (95% CI 1.119–1.615; $P = 0.002$). The stricter 2015 landmark analysis included 7,386 participants with negative assessments in 2013 and 2015 and 599 subsequent events. Its HR was attenuated to 1.255 (95% CI 0.983–1.601; $P = 0.068$), indicating that the predictive value of low baseline PEF is most pronounced within the first few years preceding a formal clinical diagnosis. The IPCW analysis censored follow-up at the first missing assessment and included 9,348 participants and 1,025 events. The resulting HR was 1.445 (95% CI 1.209–1.726; $P < 0.001$; Supplementary Table S6).

The HR was 1.448 (95% CI 1.146–1.831) in women and 1.343 (95% CI 1.071–1.684) in men (P for interaction = 0.583). Associations had the same direction among never and ever smokers and among nondrinkers and drinkers. The estimate for participants with BMI of 28 kg/m² or higher was less precise (HR 1.567, 95% CI 0.963–2.548). We found no evidence of heterogeneity by sex, smoking, alcohol consumption, or obesity (all P for interaction > 0.05 ; Fig. 4 and Supplementary Table S3).

4. Discussion

We studied 10,031 CHARLS participants without physician-diagnosed chronic lung disease or asthma at baseline. Low PEF was associated with a higher risk of first reporting physician-diagnosed chronic lung disease. Compared with non-low PEF, low PEF was associated with an approximately 41% higher adjusted hazard and a 5.1-percentage-point higher standardized 9-year risk. Estimates were similar after accounting for gaps between assessments and weighting for loss to follow-up. However, the association weakened after excluding diagnoses reported during earlier follow-up and was not statistically significant in the 2015 landmark analysis.

Our findings extend evidence that PEF may provide prognostic information beyond its established role in asthma monitoring. Previous studies have associated reduced PEF with poorer health status, greater health-care use, disability, and mortality in older adults [8, 9]. Within CHARLS, PEF below the 10th SR percentile was also associated with incident frailty [10]. Our study used individualized expected PEF and examined a distinct transition: from no reported chronic respiratory diagnosis to a newly reported physician diagnosis.

Previous CHARLS studies have linked sarcopenia, circadian syndrome, and joint trajectories of PEF and appendicular skeletal muscle mass to chronic lung disease [11–13]. Our analysis instead focused on a single baseline respiratory phenotype. Associations were similar across the primary 10th-percentile definition, the 5th-percentile definition, PEF below 80% predicted, the CAPTURE threshold, and continuous PEF SR. Thus, the main finding was not restricted to one cutoff. The spline analysis supported an overall association but provided limited evidence that its shape departed from linearity.

Several pathways may explain the association between low PEF and subsequent diagnostic reporting. Low PEF may indicate undiagnosed airflow limitation, reduced elastic recoil, small-airway dysfunction, or other pulmonary impairment already present at baseline. It may also reflect reduced respiratory muscle strength or broader physical vulnerability. Rather than invalidating the prognostic value of PEF, the attenuation observed in the 2015 landmark analysis ($P = 0.068$) provides important clinical insights. It supports the possibility that "low PEF" in these older adults largely reflects delayed diagnostic recognition of pre-existing, subclinical respiratory impairment rather than de novo disease development over a decade. In developing settings where post-bronchodilator spirometry is not routinely performed, many patients remain in a protracted undiagnosed phase despite progressively declining lung function. In this context, our findings suggest that low PEF effectively flags this hidden vulnerability, capturing a crucial "diagnostic window" to prompt clinical evaluation years before symptoms force a formal physician diagnosis. Poor expiratory performance may also accompany symptoms or functional limitations, increasing health-care contact and opportunities for diagnosis. The available CHARLS data could not distinguish among these pathways. The findings therefore should not be interpreted as evidence that low PEF causes chronic lung disease.

The clinical interpretation of these observational findings should remain cautious. Low PEF may identify adults with respiratory vulnerability who warrant further assessment, but it is not a diagnostic test for COPD. We did not assess discrimination, calibration, net benefit, or the incremental value of PEF beyond established case-finding information. The follow-up outcome was reported physician-diagnosed chronic lung disease and was not verified by medical records or post-bronchodilator spirometry. Accordingly, our findings concern subsequent diagnostic reporting and do not establish a screening threshold for clinical practice.

This study had several methodological strengths that support the robustness of its findings. CHARLS provided a large community-based cohort, four follow-up intervals, and nine years of observation. PEF quality control required at least two valid measurements and interviewer-confirmed full effort. The exposure was also standardized for its major physiological determinants. We excluded participants who reported chronic lung disease or asthma at baseline. Discrete-time survival models accommodated the interval-censored outcome, and standardized risks complemented relative estimates. Gap-corrected, landmark, IPCW, Cox, spline, alternative-threshold, and exposure-weighted analyses evaluated robustness under different assumptions. Missing covariates excluded only 1.3% of participants eligible for the PEF and outcome analyses.

Several limitations should also be considered when interpreting these findings. First, the self-reported outcome may have included chronic bronchitis, emphysema, or other chronic lung diseases. Reporting may also have depended on access to care and health-care use. Importantly, a negative baseline self-report could not definitively exclude existing mild or asymptomatic respiratory disease. While the attenuation in the 2015 landmark analysis suggests that our findings may partly reflect the delayed diagnostic recognition of pre-existing impairment (often termed reverse causation in epidemiological contexts), from a public health perspective, this limitation actually reinforces the practical utility of PEF in identifying such unrecognized patients earlier. Second, PEF is effort dependent and less specific than spirometry. Although we applied measurement-quality criteria, test posture varied, and readings at the prespecified 30- and 800-L/min boundaries were retained.

Third, the adjustment set did not include occupational exposure, air pollution, respiratory symptoms, physical activity, or health-care utilization. Residual confounding and diagnostic-opportunity bias therefore remain plausible explanations for part of the association. Linking biomarker and PEF records and restricting analyses to complete cases may also have selected a healthier or more health-engaged sample. IPCW addresses loss to observation only if the weighting models include all relevant predictors of attrition. Fourth, some scheduled disease assessments were intermittently missing during follow-up. Although gap correction indicated that interval coding did not explain the association, the diagnosis date remained unknown, and Cox analyses relied on midpoint approximations.

Finally, death may preclude a subsequent diagnosis, but the harmonized files contained only 43 deaths, an implausibly low count for this cohort and follow-up period. We therefore did not present a competing-risk analysis using these incomplete records. Survey weights and community-level clustering were not incorporated into the analysis. Thus, standardized risks describe the analytic cohort and should not be interpreted as nationally representative. The internal PEF reference equations showed modest explanatory power ($R^2 = 0.198$ in men and 0.139 in women). These equations require external validation and may not apply to other populations. Health-care use in 2020 may also have been altered by the coronavirus disease 2019 pandemic.

5. Conclusions

Low baseline PEF relative to values expected for age, sex, and height was associated with higher relative and absolute risks of subsequently reported physician-diagnosed chronic lung disease. Estimates were consistent across several sensitivity analyses but attenuated after a longer diagnostic lag. This pattern suggests that low baseline PEF largely reflects the presence of unrecognized respiratory impairment that precedes formal diagnosis by several years. Consequently, low PEF remains a highly valuable, low-cost clinical indicator that should prompt earlier evaluation for undiagnosed chronic lung diseases among middle-aged and older adults. These findings do not establish incident COPD or the clinical utility of a PEF screening threshold.

Abbreviations

BMI
body mass index
CAPTURE
COPD Assessment in Primary Care to Identify Undiagnosed Respiratory Disease and Exacerbation Risk
CHARLS
China Health and Retirement Longitudinal Study
CI
confidence interval
CLD
chronic lung disease
COPD
chronic obstructive pulmonary disease
HR
hazard ratio
IPCW
inverse probability of censoring weighting
IPTW
inverse probability of treatment weighting
PEF
peak expiratory flow
SR
standardized residual
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology

Declarations

Supplementary information

The supplementary material includes Cox models, alternative PEF definitions, subgroup analyses, IPTW diagnostics, proportional hazards tests, gap-corrected and landmark analyses, and IPCW results. It also provides internal PEF reference equations and source data for the main tables and figures. Supplementary Tables S1-S7 are cited throughout the main manuscript text.

Ethics approval and consent to participate

The original CHARLS study was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015), and all participants provided written informed consent. The present analysis used only de-identified secondary data from CHARLS.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Clinical trial number:

not applicable.

Funding

The authors received no financial support for the research, authorship, or publication of this article.

Author Contribution

Hao Haohao Chen Chen: Writing - original draft; Writing - review and editing. Yunhao Jiang: Investigation; Software; Writing - review and editing.

Data availability

The CHARLS datasets analyzed in this study are available to registered researchers through the CHARLS repository (<https://charls.pku.edu.cn/>). Derived participant-level data cannot be redistributed because access is governed by CHARLS data-use conditions. Analysis code is available from the corresponding author on reasonable request.

References

1. GBD Chronic Respiratory Disease Collaborators. Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Respir Med*. 2020;8:585–96. 10.1016/S2213-2600(20)30105-3.
2. Wang C, Xu J, Yang L, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health study): a national cross-sectional study. *Lancet*. 2018;391:1706–17. 10.1016/S0140-6736(18)30841-9.
3. Singh D, Stockley R, Anzueto A, et al. GOLD Science Committee recommendations for the use of pre- and post-bronchodilator spirometry for the diagnosis of COPD. *Eur Respir J*. 2025;65:2401603. 10.1183/13993003.01603-2024.
4. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). *Int J Epidemiol*. 2014;43:61–8. 10.1093/ije/dys203.

5. Ji C, Xia Y, Dai H, Zhao Z, Liu T, Tong S, et al. Reference values and related factors for peak expiratory flow in middle-aged and elderly Chinese. *Front Public Health*. 2021;9:706524. 10.3389/fpubh.2021.706524.
6. Vaz Fragoso CA, Gahbauer EA, Van Ness PH, Gill TM. Reporting peak expiratory flow in older persons. *J Gerontol Biol Sci Med Sci*. 2007;62:1147–51. 10.1093/gerona/62.10.1147.
7. Martinez FJ, Mannino D, Leidy NK, et al. A new approach for identifying patients with undiagnosed chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2017;195:748–56. 10.1164/rccm.201603-0622OC.
8. Roberts MH, Mapel DW. Limited lung function: impact of reduced peak expiratory flow on health status, health-care utilization, and expected survival in older adults. *Am J Epidemiol*. 2012;176:127–34. 10.1093/aje/kwr503.
9. Vaz Fragoso CA, Gahbauer EA, Van Ness PH, Concato J, Gill TM. Peak expiratory flow as a predictor of subsequent disability and death in community-living older persons. *J Am Geriatr Soc*. 2008;56:1014–20. 10.1111/j.1532-5415.2008.01687.x.
10. Wang R, Shi W, Zhou W, Xu Y, Wang J. Associations between peak expiratory flow and frailty in elderly individuals: findings from the China Health and Retirement Longitudinal Study. *Front Public Health*. 2024;12:1392581. 10.3389/fpubh.2024.1392581.
11. Wang H, Qiu H, Gu X, Zhang Y, Wang S. The association between sarcopenia and incident chronic lung disease in the general population: a longitudinal study based on CHARLS data. *Exp Gerontol*. 2023;180:112257. 10.1016/j.exger.2023.112257.
12. Liu B, Wu S, Xu Y, Jiang J, Liu Y. Prospective association between circadian syndrome and incident chronic lung disease in the CHARLS and ELSA cohorts. *Sci Rep*. 2026;16:22304. 10.1038/s41598-026-50994-1.
13. Liang J, Yang M, Tan S, Li R, Xu L, Shi Z, et al. Probable respiratory sarcopenia trajectories and the risk of incident chronic diseases: findings from the China Health and Retirement Longitudinal Study. *J Nutr Health Aging*. 2026;30:100903. 10.1016/j.jnha.2026.100903.
14. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med*. 2007;4:e296. 10.1371/journal.pmed.0040296.
15. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168:656–64. 10.1093/aje/kwn164.

Figures

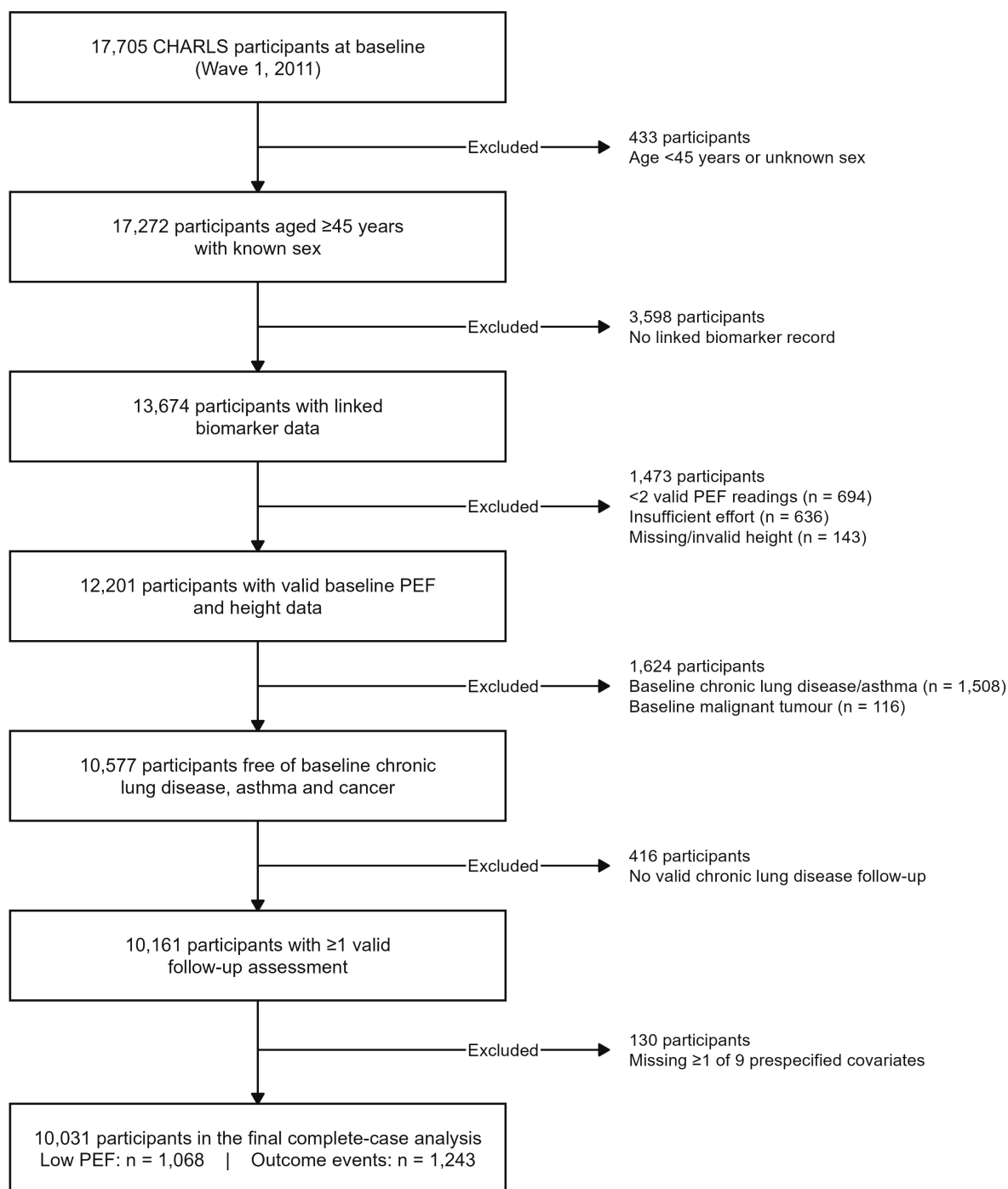


Figure 1

Participant flow through the study

The flow diagram shows construction of the 2011 baseline cohort. PEF quality control required at least two measurements between 30 and 800 L/min and interviewer-confirmed full effort. Exclusions included baseline physician-diagnosed chronic lung disease or asthma, malignant tumor, no valid follow-up

outcome, and missing prespecified covariates. CHARLS, China Health and Retirement Longitudinal Study; PEF, peak expiratory flow.

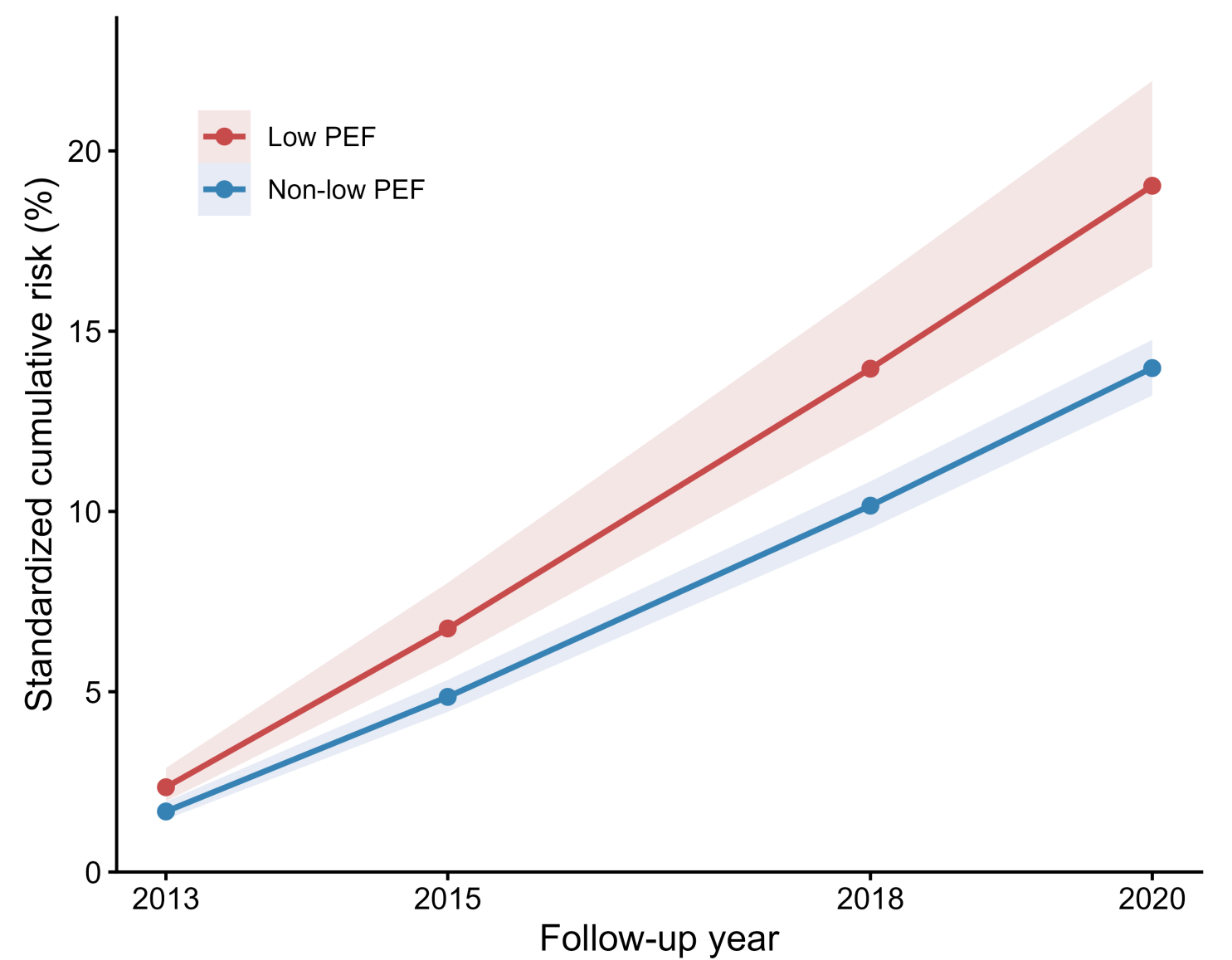


Figure 2

Adjusted standardized cumulative risk according to PEF status

Lines show model-standardized cumulative risks under low and non-low PEF at the 2013, 2015, 2018, and 2020 follow-up waves. Predictions came from the fully adjusted complementary log-log discrete-time model and were standardized to the complete cohort's covariate distribution. Shaded areas indicate model-based 95% CIs around each standardized risk curve. CI, confidence interval; PEF, peak expiratory flow.

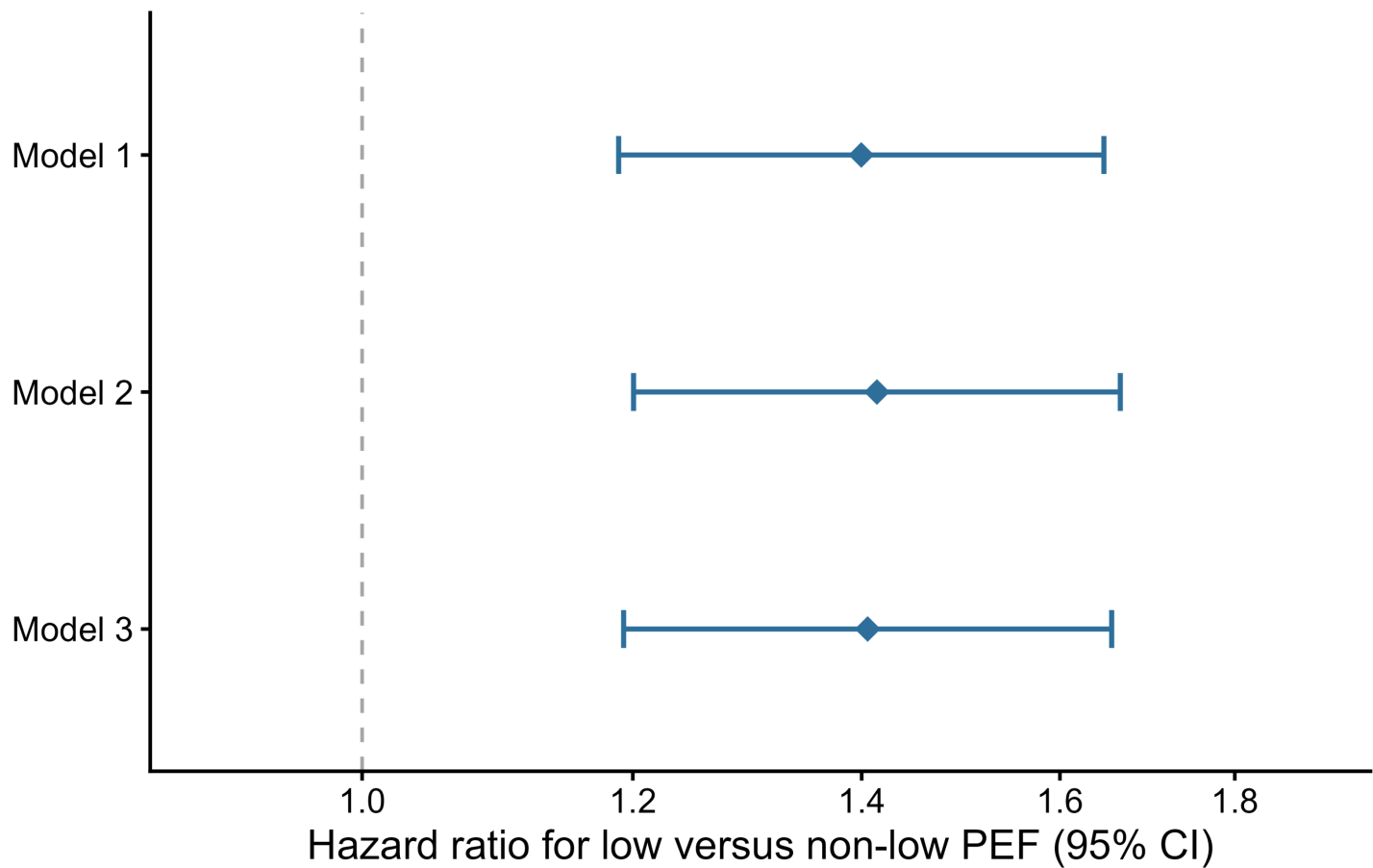


Figure 3

Association between low PEF and newly reported physician-diagnosed chronic lung disease across sequential models

Points and horizontal lines show HRs and 95% CIs for low versus non-low PEF across Models 1-3 of the complementary log-log discrete-time survival analysis. Participant-clustered robust standard errors accounted for repeated observations within individuals. All models included 10,031 participants and 1,243 events. Model definitions are provided in the Methods and Table 2. CI, confidence interval; HR, hazard ratio; PEF, peak expiratory flow.

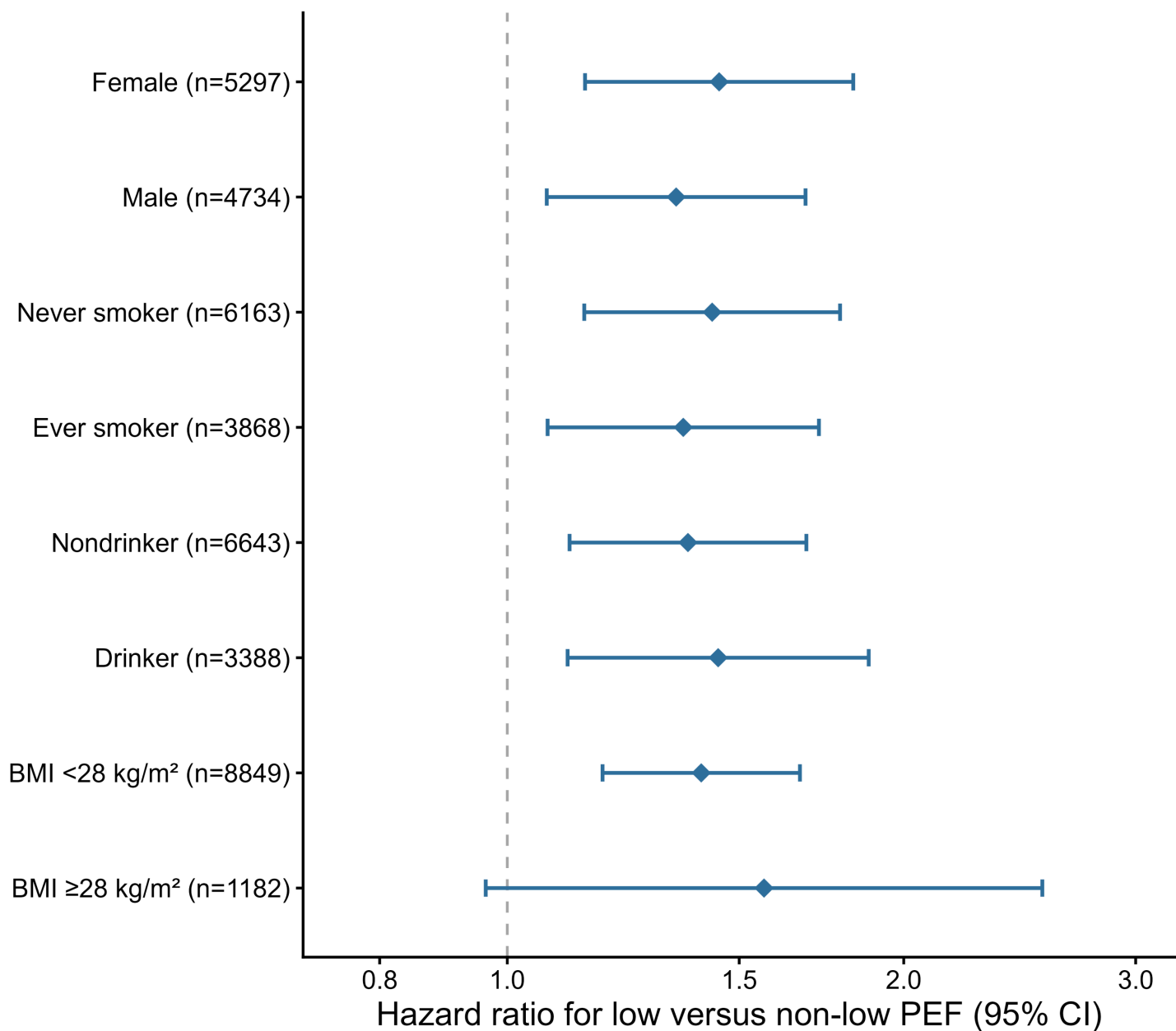


Figure 4

Subgroup associations between low PEF and newly reported physician-diagnosed chronic lung disease

Points and horizontal lines show HRs and 95% CIs from Cox proportional hazards models stratified by sex, smoking, alcohol consumption, and obesity. Models adjusted for age, sex, marital status, education, smoking, alcohol consumption, BMI, hypertension, and diabetes, excluding the relevant stratifying variable. P values for interaction were obtained from multiplicative interaction terms. BMI, body mass index; CI, confidence interval; HR, hazard ratio; PEF, peak expiratory flow.

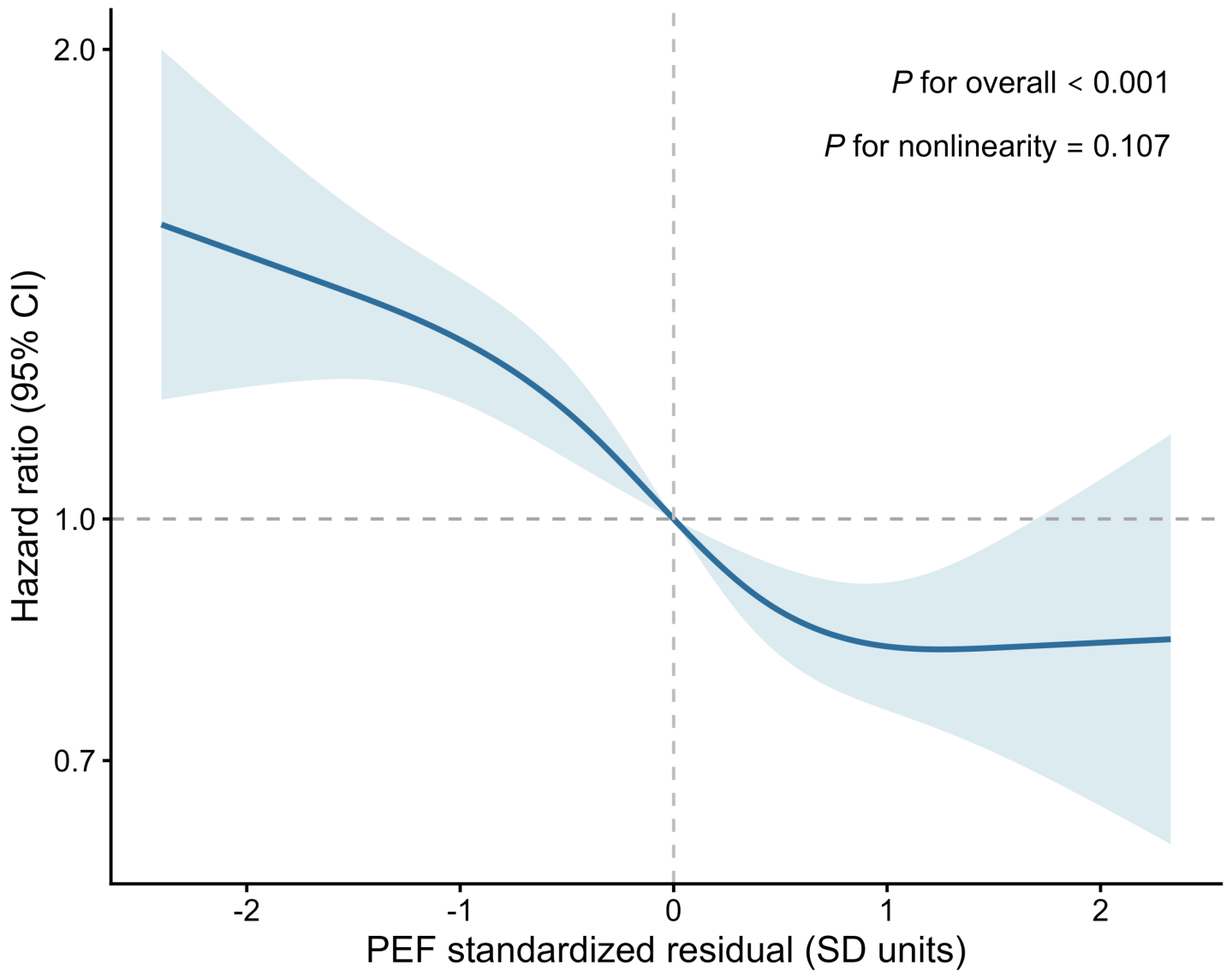


Figure 5

Dose-response association between standardized PEF and newly reported physician-diagnosed chronic lung disease

The solid line and shaded area show adjusted HRs and participant-clustered robust 95% CIs from a complementary log-log discrete-time survival model. PEF SR was modeled with a restricted cubic spline using knots at the 5th, 35th, 65th, and 95th percentiles; PEF SR 0 was the reference. The model included follow-up interval, age, sex, marital status, education, smoking, alcohol consumption, BMI, hypertension, and diabetes. Overall and nonlinearity P values were obtained from prespecified Wald tests. CI, confidence interval; HR, hazard ratio; PEF, peak expiratory flow; SR, standardized residual.