

Original Article



Associations of Genetic Risk and Physical Activity with Incident Chronic Obstructive Pulmonary Disease: A Large Prospective Cohort Study

Jin Yang^{1,&}, Xiaolin Wang^{1,&}, Wenfang Zhong^{1,&}, Jian Gao¹, Huan Chen¹, Peiliang Chen¹,
Qingmei Huang¹, Yixin Zhang¹, Fangfei You¹, Chuan Li¹, Weiqi Song¹, Dong Shen¹,
Jiaojiao Ren², Dan Liu¹, Zhihao Li¹, and Chen Mao^{1,3,#}

1. Department of Epidemiology, School of Public Health, Southern Medical University, Guangzhou 510515, Guangdong, China; 2. School of Health Management, Southern Medical University, Guangzhou 510515, Guangdong, China; 3. National Institute of Health Data Science of China, Southern Medical University, Guangzhou 510515, Guangdong, China

Abstract

Objective To investigate the relationship between physical activity and genetic risk and their combined effects on the risk of developing chronic obstructive pulmonary disease.

Methods This prospective cohort study included 318,085 biobank participants from the UK. Physical activity was assessed using the short form of the International Physical Activity Questionnaire. The participants were stratified into low-, intermediate-, and high-genetic-risk groups based on their polygenic risk scores. Multivariate Cox regression models and multiplicative interaction analyses were used.

Results During a median follow-up period of 13 years, 9,209 participants were diagnosed with chronic obstructive pulmonary disease. For low genetic risk, compared to low physical activity, the hazard ratios (HRs) for moderate and high physical activity were 0.853 (95% confidence interval [CI]: 0.748–0.972) and 0.831 (95% CI: 0.727–0.950), respectively. For intermediate genetic risk, the HRs were 0.829 (95% CI: 0.758–0.905) and 0.835 (95% CI: 0.764–0.914), respectively. For participants with high genetic risk, the HRs were 0.809 (95% CI: 0.746–0.877) and 0.818 (95% CI: 0.754–0.888), respectively. A significant interaction was observed between genetic risk and physical activity.

Conclusion Moderate or high levels of physical activity were associated with a lower risk of developing chronic obstructive pulmonary disease across all genetic risk groups, highlighting the need to tailor activity interventions for genetically susceptible individuals.

Key words: Chronic obstructive pulmonary disease; Genetic risk; Polygenic risk scores; Physical activity

Biomed Environ Sci, 2025; 38(10): 1194-1204 doi: [10.3967/bes2025.112](https://doi.org/10.3967/bes2025.112) ISSN: 0895-3988

www.besjournal.com (full text)

CN: 11-2816/Q

Copyright ©2025 by China CDC

[&]These authors contributed equally to this work.

[#]Correspondence should be addressed to Chen Mao, PhD, Tel/Fax: 86-20-61648430, E-mail: maochen9@smu.edu.cn

Biographical notes of the first authors: Jin Yang, PhD, majoring in respiratory disease, E-mail: yanaj@i.smu.edu.cn; Xiaolin Wang, MPH, majoring in chronic epidemiology, E-mail: 3235379960@qq.com; Wenfang Zhong, PhD, majoring in chronic epidemiology, E-mail: wenlyly8700@smu.edu.cn

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is the third leading cause of death globally^[1]. Between 2020 and 2050, COPD is projected to cause an economic loss of 4.326 trillion United States dollars worldwide^[2], underscoring its substantial public health burden. COPD develops as a result of interactions between genetic and environmental factors^[3,4].

The association between physical activity (PA) and lung diseases has been explored in relevant studies^[5-7]. One study showed that moderate PA enhanced lung function and improved lung capacity^[5]. Another study demonstrated that PA levels predicted important outcomes in COPD^[6], with low PA levels associated with higher mortality rates in COPD. Although PA is recommended as a preventive measure^[7], it remains uncertain whether it can effectively prevent incident COPD, particularly in individuals with varying genetic risk levels.

Although smoking is the primary risk factor for COPD, not all smokers develop the disease, suggesting that genetic factors contribute to susceptibility^[4]. More than 80 genetic loci have been linked to COPD, which influences early life events and shapes lung function over time^[8]. Advances in genome-wide association studies (GWAS) have led to the identification of multiple genetic variants associated with COPD risk and related phenotypes^[9]. Based on GWAS data, polygenic risk scores (PRS) have been developed to predict an individual's risk of COPD^[10]. One study reported that a higher PRS was associated with decreased lung function and an increased risk of COPD^[11]. However, the relationships between genetic risk scores, PA, and COPD outcomes, as well as their potential synergistic effects, remain unclear.

This study aimed to investigate the association between PA and the incidence of COPD among individuals with low, intermediate, or high genetic risk and to assess the potential interaction between genetic risk and PA in the development of COPD.

METHODS

Study Design

The UK Biobank is a comprehensive biomedical database containing de-identified genetic and health information of over 500,000 UK participants aged 40–69 years. Baseline data were collected between 2006 and 2010 using touchscreen questionnaires,

interviews, physical measurements, and health records. These data include demographic, health, and lifestyle information (such as PA), supporting vital research on common and life-threatening diseases. This prospective cohort study analyzed participants who did not have COPD at baseline and had available data on both PA and genetic risk. A total of 318,085 participants were included in the analysis (Supplementary Figure S1).

Physical Activity

PA was assessed using the short form of the International Physical Activity Questionnaire (IPAQ), which records the frequency and duration of walking, moderate-intensity, and vigorous-intensity activities within a week^[12]. The Metabolic Equivalent of Task (MET) was used to estimate the intensity of PA, with one MET defined as the energy expenditure rate of a person at rest^[13]. The IPAQ assigns MET values of 3.3 for walking, 4.0 for moderate-intensity activities, and 8.0 for vigorous-intensity activities. PA levels were calculated as MET × frequency (days/week) × duration (min/day). Total physical activity (TPA) was obtained by summing the MET values for all activity types. Based on the IPAQ recommendations, participants were categorized into three PA levels: low (< 600 MET-min/week), moderate (600–3,000 MET-min/week), and high (> 3,000 MET-min/week). The detailed guidelines can be obtained from the IPAQ data processing and analysis manual^[14].

Polygenic Risk Scores

The UK Biobank collected genetic and phenotypic data from approximately 500,000 individuals across the UK, with detailed information available from a study by Bycroft et al.^[15]. The participants were genotyped using either the custom UK Biobank Axiom array (Affymetrix) or the UK Biobank Lung Exome Variant Evaluation Axiom (Affymetrix). Genotype data were imputed using the 1000 Genomes Phase 3 and UK10K reference panels. The PRS for COPD was developed using GWAS data from individuals of European ancestry^[16]. Single nucleotide polymorphisms (SNPs) missing from the UK Biobank dataset were excluded. Independent SNPs were selected using the linkage disequilibrium clumping procedure (at $r^2 < 0.01$) based on a genome-wide significance threshold ($P < 10^{-8}$) using PLINK version 2.0 (PLINK; Broad Institute, Cambridge, MA, USA)^[17]. This approach minimizes the inclusion of highly correlated SNPs, thereby reducing multicollinearity and improving the

accuracy and statistical power of the PRS. The PRS was computed across all selected SNPs associated with COPD (Supplementary Table S1)^[18]. The PRS was calculated for each UK Biobank participant as the sum of the risk alleles (0, 1, or 2) at each locus, weighted by the natural logarithm of the estimated odds ratio (OR) for that locus^[16]. The resulting PRS values were normally distributed (Supplementary Figure S2). Quintiles of the PRS were defined using cutoffs at -0.0097 (0%), 0.0011 (20%), 0.0028 (40%), 0.0042 (60%), 0.0058 (80%), and 0.0155 (100%). The highest quintile (5th quintile) was designated as the high genetic risk group. Accordingly, PRS was categorized into three groups: low (PRS \leq 0.0011), intermediate (0.0011 < PRS \leq 0.0042), and high (PRS > 0.0042), corresponding to the lowest quintile, quintiles 2–4, and the highest quintile, respectively^[19].

Outcome

Incident COPD was identified through linkage with hospital admissions and death register records using the International Classification of Diseases (ICD) codes ICD-9 (490-492, 494, 496) and ICD-10 (J40-J44). The follow-up time was defined as the time from the date of recruitment to the earliest of the following events: the first diagnosis of COPD, censoring date (August 31, 2022, for Scotland; May 31, 2022, for Wales; October 31, 2022, for England), or date of death.

Covariates

The covariates included age, sex, race, body mass index (BMI), Townsend deprivation index (TDI), smoking status, alcohol consumption, diet, and sleep patterns. Age was categorized as < 60 years or \geq 60 years. Sex was categorized as male or female. Race was categorized as white or non-white. BMI was categorized as normal/underweight (< 25 kg/m²), overweight (25–30 kg/m²), and obese (\geq 30 kg/m²). The TDI was divided into quartiles. Smoking status was classified as current, former, or never smokers. Alcohol consumption was categorized as current, former, or never drinkers. Dietary information was collected via touchscreen questionnaires, and a healthy diet score was calculated based on the intake of vegetables, fruits, fish, unprocessed red meat, and processed meat, with scores ranging from 0 (unhealthy) to 5 (healthiest)^[20] (Supplementary Table S2). Sleep-related variables included sleep duration, insomnia, snoring, chronotype, and daytime sleepiness. Composite sleep scores ranging from 0 to 5 were calculated and categorized as

healthy sleep (\geq 4), intermediate sleep (2–3), or poor sleep (\leq 1) (Supplementary Table S3)^[21]. Detailed measurement information is available from the UK Biobank^[22].

Statistical Analysis

Missing covariate values were imputed using multiple imputations by chained equations with five imputations (R package 'mice')^[23]. The detailed missing rates are provided in Supplementary Table S4. Cox proportional hazard regression models were used to estimate the hazard ratios (HRs) and 95% confidence intervals (CIs) for the independent and combined associations between PA, genetic risk, and incident COPD after adjusting for age, sex, BMI, race, TDI, smoking status, drinking status, sleep scores, and diet scores. The proportional hazard assumption was tested using Schoenfeld residuals, and no violation of this assumption was observed. Genetic data analyses were additionally adjusted for the genotyping chip and the first ten principal components of ancestry.

The interactions between PRS and PA were tested by adding a cross-product term to the Cox model. The statistical significance of the interaction was assessed using the likelihood ratio test by comparing the goodness of fit between models with and without the interaction term. Nonlinear relationships among genetic risk, TPA, and incident COPD were evaluated using restricted cubic spline curves in multivariable Cox models, with four knots placed at the 5th, 35th, 65th, and 95th percentiles of TPA (referred to as the 5th percentile).

Sensitivity and subgroup analyses were performed. First, to minimize reverse causality, participants who developed COPD within the first two years of follow-up were excluded from the analysis. Second, we reanalyzed the risk of incident COPD by removing missing data from all variables. Third, Cox proportional hazards models were conducted stratified by age groups (\leq 45, 45–55, 55–65, > 65 years) and sex (male or female). To account for multiple testing across models and covariates, *P*-values were adjusted using the Benjamini-Hochberg method to control for the false discovery rate (FDR). Statistical significance was determined at an FDR-adjusted *P*-value of < 0.05. Finally, to mitigate potential selection and confounding biases, we used multivariable-adjusted propensity score matching (PSM) to assess the robustness of the results^[24] after adjusting for age, sex, BMI, race, TDI, smoking status, drinking status, sleep scores, diet scores, and other covariates. A

matching ratio of 1:n should not exceed 1:4 to avoid overfitting. In this study, we used a 1:1 to 1:3 PSM (matching low PA with moderate-to-high PA in a ratio of 1:1–1:3).

All analyses were conducted using the R software (version 4.3.1, R Project for Statistical Computing, Vienna, Austria). All statistical tests were two-sided, with $P < 0.05$ considered statistically significant.

RESULTS

Baseline Characteristics

Among the 318,085 participants, 9,209 were diagnosed with COPD during a median follow-up period of 13.0 years. Of these participants, 131,931 (41.5%) were aged 60 years or older, 169,135 (53.2%) were female, and 303,205 (95.3%) were White. With regard to genetic risk, 63,678 participants (20.0%) had low genetic risk, 127,230 (40.0%) had intermediate genetic risk, and 127,177 (40.0%) had high genetic risk. Most participants achieved moderate or high PA levels (> 80.0%). Compared with participants without COPD, those who developed COPD were significantly more likely to be aged 60 years or older, male, obese, in the fourth quartile of the TDI, current or former smokers, former drinkers, have a healthy diet score of < 3, and have poor sleep scores (Table 1).

Association between Physical Activity, Genetic Risk, and Incident COPD

Model 1 was adjusted for age, sex, BMI, race, TDI, smoking status, drinking status, sleep score, and dietary score (Table 2). Compared with low PA, moderate and high PA were associated with a lower risk of developing COPD, with HRs of 0.822 (95% CI : 0.779–0.868) and 0.824 (95% CI : 0.780–0.871), respectively. Compared with low genetic risk, intermediate and high genetic risks were associated with a higher risk of developing COPD, with HRs of

1.105 (95% CI : 1.042–1.173) and 1.339 (95% CI : 1.264–1.418), respectively. Model 2 was further adjusted for either the genetic risk category or PA level, and the results remained consistent. Compared with low PA, moderate and high PA were associated with HRs of 0.824 (95% CI : 0.780–0.870) and 0.826 (95% CI : 0.782–0.873), respectively. Compared with low genetic risk, intermediate and high genetic risks were associated with HRs of 1.105 (95% CI : 1.041–1.172) and 1.337 (95% CI : 1.262–1.416), respectively.

Impact of Physical Activity on Incident COPD across Different Genetic Risk

A significant interaction was observed between PA and genetic risk for incident COPD (P for interaction < 0.001). Based on this, we conducted a stratified analysis by genetic risk and observed that moderate and high PA had protective effects on incident COPD across different genetic risk categories compared to low PA. For individuals with low genetic risk, compared with low PA, the HRs for moderate and high PA were 0.853 (95% CI : 0.748–0.972) and 0.831 (95% CI : 0.727–0.950), respectively. Among those with intermediate genetic risk, the HRs for moderate and high PA were 0.829 (95% CI : 0.758–0.905) and 0.835 (95% CI : 0.764–0.914), respectively. In the high genetic risk group, the HRs were 0.809 (95% CI : 0.746–0.877) and 0.818 (95% CI : 0.754–0.888), respectively (Figure 1). In summary, compared with low PA, moderate and high PA were associated with an approximately 20% lower risk of COPD. We further examined the combined effects of PA levels and genetic susceptibility. Individuals with a low genetic risk and high PA had a 40% lower risk of developing COPD than those with a high genetic risk and low PA (Figure 2). A dose-response relationship was observed, with the lowest risk of developing COPD observed within the range of 1,800–1,900 MET-min/week across all genetic risk

Table 1. Baseline characteristics of the study participants stratified by COPD

Characteristic	Overall (n = 318,085)	No Incident COPD (n = 308,876)	Incident COPD (n = 9,209)	P value
Age, n (%)				< 0.001
< 60 years	186,154 (58.5)	183,236 (59.3)	2,918 (31.7)	
≥ 60 years	131,931 (41.5)	125,640 (40.7)	6,291 (68.3)	
Sex, n (%)				< 0.001
Male	148,950 (46.8)	143,729 (46.5)	5,221 (56.7)	
Female	169,135 (53.2)	165,147 (53.5)	3,988 (43.3)	

Continued

Characteristic	Overall (n = 318,085)	No Incident COPD (n = 308,876)	Incident COPD (n = 9,209)	P value
Race, n (%)				< 0.001
White	303,205 (95.3)	294,256 (95.3)	8,949 (97.2)	
Non-white	14,880 (4.7)	14,620 (4.73)	260 (2.82)	
BMI, n (%)				< 0.001
Normal/underweight ($< 25 \text{ kg/m}^2$)	108,030 (34.0)	105,440 (34.1)	2,590 (28.1)	
Overweight (25–30 kg/m^2)	136,794 (43.0)	133,122 (43.1)	3,672 (39.9)	
Obese ($\geq 30 \text{ kg/m}^2$)	73,261 (23.0)	70,314 (22.8)	2,947 (32.0)	
TDI, mean (SD), n (%)				< 0.001
Quartile 1	79,527 (25.0)	77,948 (25.2)	1,579 (17.1)	
Quartile 2	79,520 (25.0)	77,708 (25.2)	1,812 (19.7)	
Quartile 3	79,520 (25.0)	77,339 (25.0)	2,181 (23.7)	
Quartile 4	79,518 (25.0)	75,881 (24.6)	3,637 (39.5)	
Smoking status, n (%)				< 0.001
Current	30,839 (9.7)	27,618 (8.94)	3,221 (35.0)	
Former	111,498 (35.1)	107,261 (34.7)	4,237 (46.0)	
Never	175,748 (55.3)	173,997 (56.3)	1,751 (19.0)	
Drinking status, n (%)				< 0.001
Current	296,005 (93.1)	287,750 (93.2)	8,255 (89.6)	
Former	10,127 (3.2)	9,540 (3.09)	587 (6.37)	
Never	11,953 (3.8)	11,586 (3.75)	367 (3.99)	
Diet scores, n (%)				< 0.001
0	5,171 (1.6)	4,897 (1.59)	274 (2.98)	
1	34,219 (10.8)	32,866 (10.6)	1,353 (14.7)	
2	82,112 (25.8)	79,429 (25.7)	2,683 (29.1)	
3	107,403 (33.8)	104,642 (33.9)	2,761 (30.0)	
4	67,989 (21.4)	66,373 (21.5)	1,616 (17.5)	
5	21,191 (6.7)	20,669 (6.69)	522 (5.67)	
Sleep scores, n (%)				< 0.001
Poor	28,098 (8.8)	26,730 (8.65)	1,368 (14.9)	
Intermediate	185,614 (58.4)	179,834 (58.2)	5,780 (62.8)	
Healthy	104,373 (32.8)	102,312 (33.1)	2,061 (22.4)	
Physical activity level, n (%)				< 0.001
Low	58,370 (18.4)	56,277 (18.2)	2,093 (22.7)	
Moderate	130,218 (40.9)	126,624 (41.0)	3,594 (39.0)	
High	129,497 (40.7)	125,975 (40.8)	3,522 (38.2)	
PRS, n (%)				< 0.001
Low	63,678 (20.0)	62,071 (20.1)	1,607 (17.5)	
Intermediate	127,230 (40.0)	123,750 (40.1)	3,480 (37.8)	
High	127,177 (40.0)	123,055 (39.8)	4,122 (44.8)	

Note. COPD, chronic obstructive pulmonary disease; BMI, body mass index; TDI, Townsend deprivation index; PA, physical activity; PRS, polygenic risk scores. All statistical tests were two-sided.

strata (Figure 3).

Subgroup Analyses

To better understand these associations across different population subgroups, we performed stratified analyses according to age and sex. For

individuals with high genetic risk, the *HRs* for moderate PA were 0.720 (95% *CI*: 0.591–0.879), 0.842 (95% *CI*: 0.753–0.941), and 0.775 (95% *CI*: 0.663–0.905) in the 45–55, 55–65, and > 65-year age groups, respectively. For participants with high PA, the *HRs* in the 55–65 and > 65 years age groups were

Table 2. Risk of incident COPD according to physical activity level or genetic risk

Subgroup	Cases/ <i>N</i>	Model 1		Model 2	
		<i>HR</i>	<i>P</i> value	<i>HR</i>	<i>P</i> value
PA level					
Low	2,093/58,370	1.000 (reference)	–	1.000 (reference)	–
Moderate	3,594/130,218	0.822 (0.779–0.868)	< 0.001	0.824 (0.780–0.870)	< 0.001
High	3,522/129,497	0.824 (0.780–0.871)	< 0.001	0.826 (0.782–0.873)	< 0.001
PRS					
Low	1,607/63,678	1.000 (reference)	–	1.000 (reference)	–
Intermediate	3,480/127,230	1.105 (1.042–1.173)	< 0.001	1.105 (1.041–1.172)	< 0.001
High	4,122/127,177	1.339 (1.264–1.418)	< 0.001	1.337 (1.262–1.416)	< 0.001

Note. Data are expressed as *HR* (95% *CI*); Model 1 was adjusted for age, sex, BMI, race, TDI, smoking status, drinking status, sleep score, and diet score. Model 2 was additionally adjusted for genetic risk category and physical activity level. All statistical tests were two-sided. *HR*, hazard ratio; *CI*, confidence interval; PA, physical activity; PRS, polygenic risk score; COPD, chronic obstructive pulmonary disease; BMI, body mass index; TDI, Townsend deprivation index.

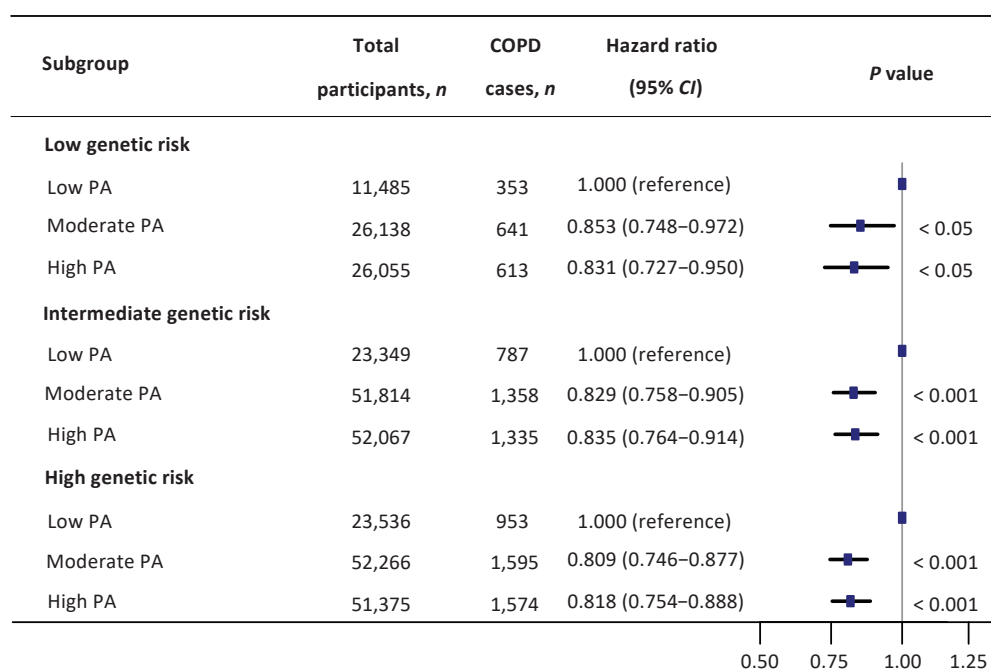


Figure 1. Combined effect of genetic risk and physical activity on incident COPD. The multivariate model was adjusted for age, sex, BMI, race, TDI, smoking status, drinking status, sleep scores, diet scores, genetic risk categories, and low physical activity levels. All statistical tests were two-sided. *HR*, hazard ratio; *CI*, confidence interval; PA, physical activity; COPD, chronic obstructive pulmonary disease; BMI, body mass index; TDI, Townsend deprivation index.

0.846 (95% *CI*: 0.764–0.959) and 0.727 (95% *CI*: 0.620–0.853), respectively (Supplementary Table S5). Among females with low genetic risk, compared to those with low PA, the *HRs* for moderate and high PA were 0.756 (95% *CI*: 0.622–0.919) and 0.749 (95% *CI*: 0.613–0.915), respectively. Among females with intermediate genetic risk, the *HRs* were 0.874 (95% *CI*: 0.766–0.997) and 0.750 (95% *CI*: 0.652–0.862), respectively. Among females with high genetic risk, moderate and high PA levels were associated with *HRs* of 0.736 (95% *CI*: 0.652–0.831) and 0.763 (95% *CI*: 0.673–0.865), respectively. While moderate and high PA levels had protective effects against incident COPD in both sexes, the effect was more pronounced in females (Supplementary Table S6).

Sensitivity Analyses

We conducted several sensitivity analyses. First, the exclusion of participants with missing covariates yielded results consistent with the main findings (Supplementary Table S7). To minimize reverse causality, participants who developed COPD within the first two years of follow-up were excluded, and the associations remained stable (Supplementary Table S8). Finally, the PSM approach was used to rebalance the distribution of the low and moderate-to-high PA groups at a 1:1 to 1:3 ratio, confirming the robustness of the findings (Supplementary

Tables S9–S12).

DISCUSSION

In this large prospective cohort study involving 318,085 individuals without a prior diagnosis of COPD, we observed that PA had a protective effect against the development of COPD across different genetic risk groups. A U-shaped dose-response relationship was observed between TPA and incident COPD, with the greatest protective effect occurring within the range of 1,800–1,900 MET-min/week. Among individuals with varying levels of genetic risk, moderate or high PA was associated with an approximately 20% lower risk of developing COPD than was low PA. These results highlight the advantages of PA in preventing COPD.

Previous studies have explored the impact of PRS on COPD or lung function as well as the role of PA as a protective factor against COPD risk^[25–28]. For instance, a GWAS involving over 400,000 individuals identified 139 new lung function-related signals that predicted COPD across multiethnic cohorts^[25]. Other studies have reported an inverse association between PA and COPD risk, indicating PA as an effective primary prevention strategy^[26–28]. However, few studies have examined the combined effects of PA and genetic susceptibility on the incidence of

Subgroup	Total participants, <i>n</i>	COPD cases, <i>n</i>	Hazard ratio (95% <i>CI</i>)	<i>P</i> value
High genetic risk + Low PA	23,536	953	1.000 (reference)	
High genetic risk + Moderate PA	52,266	1,595	0.811 (0.748–0.879)	< 0.001
High genetic risk + High PA	51,375	1,574	0.822 (0.758–0.892)	< 0.001
Intermediate genetic risk + Low PA	23,349	787	0.817 (0.744–0.898)	< 0.001
Intermediate genetic risk + Moderate PA	51,814	1,358	0.675 (0.621–0.734)	< 0.001
Intermediate genetic risk + High PA	52,067	1,335	0.679 (0.624–0.738)	< 0.001
Low genetic risk + Low PA	11,485	353	0.731 (0.647–0.826)	< 0.001
Low genetic risk + Moderate PA	26,138	641	0.623 (0.564–0.689)	< 0.001
Low genetic risk + High PA	26,055	613	0.606 (0.547–0.671)	< 0.001

Figure 2. Joint effect of physical activity and genetic susceptibility on incident COPD. The multivariate model was adjusted for age, sex, BMI, race, TDI, smoking status, drinking status, sleep scores, diet scores, genetic risk categories, and physical activity levels. All statistical tests were two-sided; *HR*, hazard ratio; *CI*, confidence interval; PA, physical activity; COPD, chronic obstructive pulmonary disease; BMI, body mass index; TDI, Townsend deprivation index.

COPD. Most existing studies have focused on patients already diagnosed with COPD^[29,30], leaving it unclear whether different PA levels can mitigate COPD risk in individuals with different genetic profiles. Genetic susceptibility to COPD varies among individuals and is believed to play a substantial role. GWAS evaluating common genetic polymorphisms, such as SNPs, have identified many genetic variants associated with COPD^[31]. To capture the cumulative genetic influence, we constructed a genetic risk score based on multiple SNPs. In our study, individuals with moderate or high levels of PA had a lower risk of incident COPD compared to those with low PA (moderate vs. low, $HR = 0.822$, $P < 0.001$; high vs. low, $HR = 0.824$, $P < 0.001$). Moreover, across all levels of genetic risk, moderate or high levels of PA were consistently associated with an approximately 15%–20% lower risk of COPD. These findings underscore the importance of PA as a

modifiable behavior that can mitigate genetic predisposition to COPD. Our conclusions align with a previous study^[32]. PA may reduce systemic inflammation by lowering levels of pro-inflammatory markers such as interleukin-6 (IL-6), tumor necrosis factor alpha (TNF- α), and C-reactive protein while upregulating anti-inflammatory mediators, including interleukin-10 (IL-10) and transforming growth factor beta (TGF- β). This anti-inflammatory effect may help slow the decline in lung function^[5]. In addition, PA increases the number of CD4-positive T cells and enhances immune function, potentially explaining the observed reduction^[33]. Regular PA influences epigenetic regulation, including the DNA methylation of genes involved in the inflammatory response, oxidative stress, and immune response^[34], mechanisms closely linked to the pathogenesis of COPD^[35].

COPD is the leading cause of morbidity,

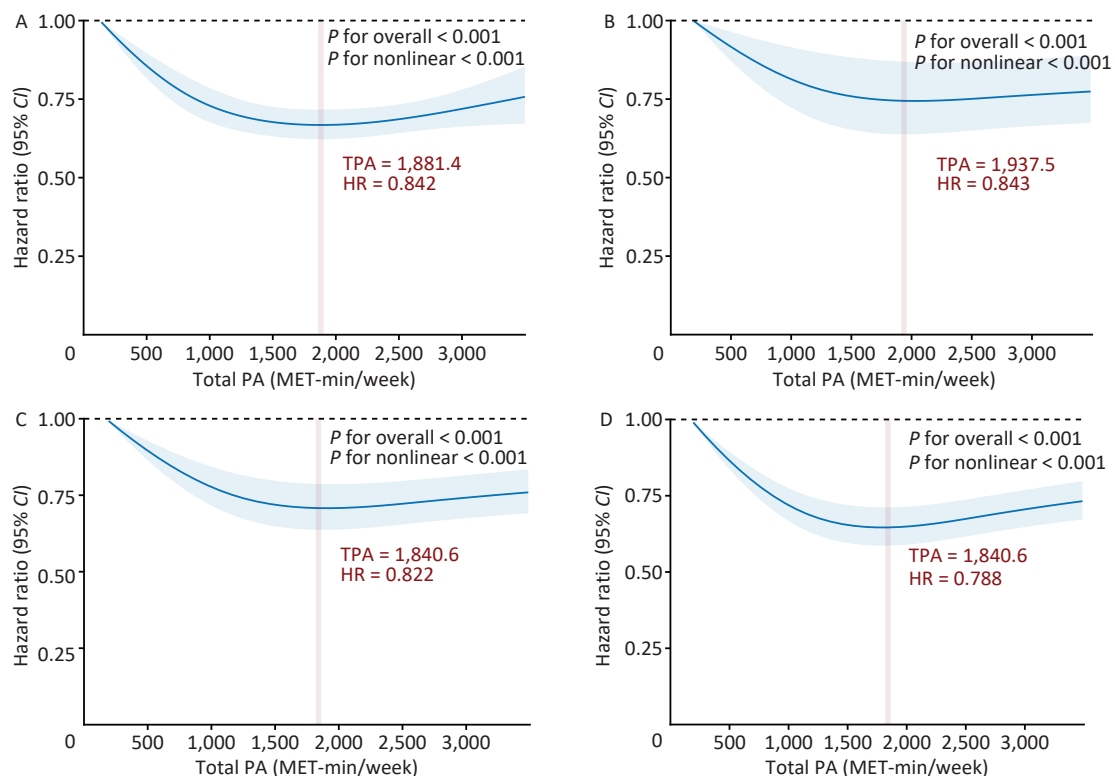


Figure 3. Dose-response associations between TPA and incident COPD. (A–D) Dose-response associations between TPA and incident COPD in (A) the general population, (B) low genetic risk, (C) moderate genetic risk, and (D) high genetic risk groups. Bold lines represent HRs, and shaded areas indicate 95% CI. Red font indicates the threshold of the restricted cubic spline curve. All data were adjusted for age, sex, BMI, race, TDI, smoking status, drinking status, sleep scores, diet scores, and genetic risk categories. All statistical tests were two-sided. HR, hazard ratio; CI, confidence interval; MET, metabolic equivalent of task; COPD, chronic obstructive pulmonary disease; PA, physical activity; BMI, body mass index; TDI, Townsend deprivation index.

mortality, and healthcare burden globally^[36]. Its pathogenesis involves a complex interaction between the genetic background and environmental factors. PA is crucial for maintaining functional status and is linked to reduced risks of adverse outcomes, such as all-cause mortality^[37], cancer^[38], cardiovascular disease^[39], and depression^[40]. Regular PA was significantly correlated with improved health outcomes, enhanced PA capacity, quality of life, and delayed disease progression. Research indicated that PA positively affects individuals with COPD^[41]. However, in real-world settings, sustaining PA among patients with COPD is challenging because of symptoms such as dyspnea, which limits PA levels^[42]. From a public health perspective, preventing COPD by promoting timely PA appears to be a promising strategy. Our results demonstrate a significant inverse association between moderate or high PA and the risk of COPD onset, even in those with elevated genetic susceptibility. The observed U-shaped dose-response relationship further suggests an optimal PA range of 1,800–1,900 MET-min/week for maximal benefit. These results are consistent with those of previous studies. One study reported a U-shaped relationship between PA and COPD risk, with moderate PA (approximately 1,500–2,000 MET-min/week) associated with the lowest COPD risk, whereas PA levels below or above this range were associated with an increased risk^[32]. Another study noted a nonlinear relationship between PA, mortality, and hospitalization rates associated with lower respiratory tract infections, with the lowest risk observed at a PA level of 1,500–1,800 MET-min/week and a decreasing trend in risk beyond this range^[43]. The U-shaped relationship between PA and the risk of COPD may reflect the dual effects of PA on lung health. Moderate PA helps enhance lung function, reduce inflammation levels, and improve immune function, thereby lowering the risk of COPD. However, extremely low or excessively high PA levels may have adverse effects, potentially increasing the risk of COPD. Excessive PA can contribute to the development or progression of COPD through mechanisms such as oxidative stress, inflammation, mechanical stress, or lung injury. Therefore, it is recommended that individuals engage in a moderate amount of PA, specifically approximately 1,800–1,900 MET-min/week, to achieve optimal lung health benefits^[44].

COPD is closely related to age, with lung function typically declining with age, making older adults more susceptible to the disease. Genetic factors may contribute to COPD development in younger

individuals^[45]. The pathogenesis of COPD may begin earlier in life, and respiratory disorders during childhood and adolescence can increase the risk of reduced lung function in adulthood^[46]. Moreover, early-onset COPD does not necessarily indicate a mild disease. One study showed that the decline in lung function is much faster in the early stages than in the later stages^[47] and that patients with clinical COPD under 55 years of age may experience disease severity similar to those over 65 years^[48]. Therefore, early intervention and health promotion measures can significantly reduce the risk and severity of this disease. Our research observed that in the 45–55 year age group, regardless of low, intermediate, or high genetic risk, moderate or high levels of PA were associated with a reduced risk of COPD compared with low PA, and these individuals may benefit more than those aged 55 years and above. Sex-stratified subgroup analyses indicated that varying levels of PA were linked to a 7%–21% decreased risk of incident COPD in males and a 13%–27% decreased risk in females, depending on the genetic risk level. Individuals of different sexes may differ in their lifestyles and behavior patterns^[49–51]. One study observed that females achieved the greatest survival benefit with approximately 140 min of moderate-to-vigorous-intensity PA per week, whereas males required approximately 300 min to achieve similar benefits^[52]. Additionally, females are more proactive in seeking healthcare and may be more consistent in their health-promoting behaviors, including regular PA. This increased health awareness could lead to earlier detection and management of COPD symptoms, potentially enhancing the protective effects of PA^[53]. These studies support our results.

This study has several strengths, including its prospective design, large sample size, more than ten years of follow-up, and comprehensive PA data. A unique feature of this study is its investigation of the combined influence of PA and genetic risk on the incidence of COPD.

However, this study has certain limitations. First, as an observational study, we could not establish a causal relationship between PA and incident COPD, although our results remained largely consistent after excluding participants who developed COPD within two years. Second, the age range of participants was limited to 40–69 years, which may not represent the broader population. Third, PA was assessed *via* self-reporting, which may have caused information bias. Future studies should consider combining accelerometer-based measurements with self-report methods to validate our findings. Fourth,

the study population was limited to individuals of European ancestry, which potentially limits its generalizability. Genetic and lifestyle differences across populations may have influenced the observed associations. Future studies should include more diverse cohorts. Finally, patients with COPD were identified through hospital records and death registries, which may have missed a few patients with COPD who did not seek medical care or were not hospitalized or may have missed mild or undiagnosed cases.

CONCLUSION

Compared to low PA, moderate or high PA was significantly associated with a lower risk of COPD across the different genetic risk groups. The dose-response relationship showed a U-shaped association between TPA and incident COPD. Individuals with a low genetic risk and high PA levels exhibited the lowest risk, whereas those with a high genetic risk benefitted significantly from moderate and high PA levels. These findings highlight the potential of PA intervention as an effective strategy to mitigate the risk of COPD across the entire population, including those with a high genetic predisposition to the disease.

Funding This work was supported by the Construction of High-level University of Guangdong (G624330242), the National Natural Science Foundation of China (82425052) to Dr. Chen Mao, and the Postdoctoral Fellowship Program of CPSF (GZC20231052) to Dr. Jin Yang.

Competing Interests The authors declare that they have no competing interests.

Ethics This study was approved by the Northwest Multicenter Research Ethics Committee (11/NW/0382) and was conducted using the UK Biobank Resource (application number 43795). All the participants provided written informed consent.

Authors' Contributions Conception and design of the research: Chen Mao, Jin Yang. Data curation: Jin Yang, Wenfang Zhong. Data analysis and manuscript drafting: Jin Yang, Xiaolin Wang. Manuscript editing and revision: Jin Yang, Xiaolin Wang, Wenfang Zhong, Jian Gao, Huan Chen, Peiliang Chen, Qingmei Huang, Yixin Zhang, Fangfei You, Chuan Li, Weiqi Song, Dong Shen, Jiaojiao Ren, Dan Liu, Zhihao Li, Chen Mao. All authors approved the final version of the manuscript.

Acknowledgements The authors thank the UK Biobank.

Data Sharing The UK Biobank data are available upon request (www.ukbiobank.ac.uk/). The supplementary materials will be available in www.besjournal.com.

Received: January 6, 2025;

Accepted: June 10, 2025

REFERENCES

1. World Health Organization. The top 10 causes of death. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>. [2025-04-15]
2. Chen SM, Kuhn M, Prettner K, et al. The global economic burden of chronic obstructive pulmonary disease for 204 countries and territories in 2020-50: a health-augmented macroeconomic modelling study. *Lancet Global Health*, 2023; 11, e1183-93.
3. Global Initiative for Chronic Obstructive Lung Disease. 2024 gold report - global initiative for chronic obstructive lung disease - GOLD. <https://goldcopd.org/2024-gold-report/>. [2024-06-21]
4. Agustí A, Melén E, DeMeo DL, et al. Pathogenesis of chronic obstructive pulmonary disease: understanding the contributions of gene-environment interactions across the lifespan. *Lancet Respir Med*, 2022; 10, 512-24.
5. Garcia-Aymerich J, Lange P, Benet M, et al. Regular physical activity modifies smoking-related lung function decline and reduces risk of chronic obstructive pulmonary disease: a population-based cohort study. *Am J Respir Crit Care Med*, 2007; 175, 458-63.
6. Spruit MA, Pitta F, McAuley E, et al. Pulmonary rehabilitation and physical activity in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 2015; 192, 924-33.
7. Perry AS, Dooley EE, Master H, et al. Physical activity over the lifecourse and cardiovascular disease. *Circ Res*, 2023; 132, 1725-40.
8. Sakornsakolpat P, Prokopenko D, Lamontagne M, et al. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat Genet*, 2019; 51, 494-505.
9. Cho MH, McDonald MLN, Zhou XB, et al. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med*, 2014; 2, 214-25.
10. Zhang PD, Zhang XR, Zhang A, et al. Associations of genetic risk and smoking with incident COPD. *Eur Respir J*, 2022; 59, 2101320.
11. Hernandez-Pacheco N, Kilanowski A, Kumar A, et al. Exploring the genetics of airflow limitation in lung function across the lifespan - a polygenic risk score study. *eClinicalMedicine*, 2024; 75, 102731.
12. Bassett DR Jr. Commentary to accompany: international physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*, 2003; 35, 1396.
13. IPAQ. Guidelines for data processing and analysis of the international physical activity questionnaire (IPAQ). https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/ipaq_analysis.pdf. [2025-04-15]
14. IPAQ. Guidelines for data processing and analysis of the international physical activity questionnaire (IPAQ). <https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=540>. [2024-07-03]

15. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*, 2018; 562, 203–9.
16. Moll M, Sakornsakolpat P, Shrine N, et al. Chronic obstructive pulmonary disease and related phenotypes: polygenic risk scores in population-based and case-control cohorts. *Lancet Respir Med*, 2020; 8, 696–708.
17. PLINK 2.0 alpha. <https://www.cog-genomics.org/plink/2.0/>. [2025-04-15]
18. Zhang JZ, Xu HF, Qiao DD, et al. A polygenic risk score and age of diagnosis of COPD. *Eur Respir J*, 2022; 60, 2101954.
19. O'Sullivan JW, Ashley EA, Elliott PM. Polygenic risk scores for the prediction of cardiometabolic disease. *Eur Heart J*, 2023; 44, 89–99.
20. Feng HL, Yang LL, Liang YY, et al. Associations of timing of physical activity with all-cause and cause-specific mortality in a prospective cohort study. *Nat Commun*, 2023; 14, 930.
21. Fan MY, Sun DJY, Zhou T, et al. Sleep patterns, genetic susceptibility, and incident cardiovascular disease: a prospective study of 385 292 UK biobank participants. *Eur Heart J*, 2020; 41, 1182–9.
22. Biobank. UK biobank: browse by category. https://biobank.ndph.ox.ac.uk/showcase/browse.cgi?id=22032&cd=data_field. [2024-07-06].
23. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res*, 1999; 8, 3–15.
24. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Res*, 2011; 46, 399–424.
25. Shrine N, Guyatt AL, Erzurumluoglu AM, et al. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet*, 2019; 51, 481–93.
26. Lee YH, Huang YP, Pan SL. Physical activity and the risk of chronic obstructive pulmonary disease: a longitudinal follow-up study in Taiwan. *Am J Med Sci*, 2024; 368, 55–60.
27. Chen L, Cai M, Li HT, et al. Risk/benefit tradeoff of habitual physical activity and air pollution on chronic pulmonary obstructive disease: findings from a large prospective cohort study. *BMC Med*, 2022; 20, 70.
28. Kim T, Kim H, Kong SG, et al. Association between regular moderate to vigorous physical activity initiation following COPD diagnosis and mortality: an emulated target trial using nationwide cohort data. *Chest*, 2024; 165, 84–94.
29. Burtin C, Mohan D, Troosters T, et al. Objectively measured physical activity as a COPD clinical trial outcome. *Chest*, 2021; 160, 2080–100.
30. Cordova-Rivera L, Gibson PG, Gardiner PA, et al. Physical activity associates with disease characteristics of severe asthma, bronchiectasis and COPD. *Respirology*, 2019; 24, 352–60.
31. Stolz D. Chronic obstructive pulmonary disease risk: does genetics hold the answer? *Lancet Respir Med*, 2020; 8, 653–54.
32. Yang QL, Ma TY, Cheung K, et al. Wearable device-measured bouts and sporadic physical activity patterns and incident COPD: a prospective cohort study from UK Biobank. *Public Health*, 2024; 237, 322–29.
33. AJMC. Exercise rehabilitation found to help alleviate symptoms of several chronic respiratory diseases. <https://www.ajmc.com/view/exercise-rehabilitation-found-to-help-alleviate-symptoms-of-several-chronic-respiratory-diseases>. [2025-04-15]
34. Plaza-Diaz J, Izquierdo D, Torres-Martos Á, et al. Impact of physical activity and exercise on the epigenome in skeletal muscle and effects on systemic metabolism. *Biomedicines*, 2022; 10, 126.
35. Kaur G, Batra S. Regulation of DNA methylation signatures on NF-κB and STAT3 pathway genes and TET activity in cigarette smoke extract-challenged cells/COPD exacerbation model *in vitro*. *Cell Biol Toxicol*, 2020; 36, 459–80.
36. Christenson SA, Smith BM, Bafadhel M, et al. Chronic obstructive pulmonary disease. *Lancet*, 2022; 399, 2227–42.
37. Celis-Morales CA, Lyall DM, Anderson J, et al. The association between physical activity and risk of mortality is modulated by grip strength and cardiorespiratory fitness: evidence from 498 135 UK-Biobank participants. *Eur Heart J*, 2017; 38, 116–22.
38. Stamatakis E, Ahmadi MN, Friedenreich CM, et al. Vigorous intermittent lifestyle physical activity and cancer incidence among nonexercising adults: the UK Biobank accelerometry study. *JAMA Oncol*, 2023; 9, 1255–9.
39. Khurshid S, Al-Alusi MA, Churchill TW, et al. Accelerometer-derived "weekend warrior" physical activity and incident cardiovascular disease. *JAMA*, 2023; 330, 247–52.
40. Choi KW, Chen CY, Stein MB, et al. Assessment of bidirectional relationships between physical activity and depression among adults: a 2-sample mendelian randomization study. *JAMA Psychiatry*, 2019; 76, 399–408.
41. Xiang XY, Huang LH, Fang Y, et al. Physical activity and chronic obstructive pulmonary disease: a scoping review. *BMC Pulm Med*, 2022; 22, 301.
42. Watz H, Pitta F, Rochester CL, et al. An official European respiratory society statement on physical activity in COPD. *Eur Respir J*, 2014; 44, 1521–37.
43. Jung MH, Yi SW, An SJ, et al. Association of physical activity and lower respiratory tract infection outcomes in patients with cardiovascular disease. *J Am Heart Assoc*, 2022; 11, e023775.
44. American Lung Association. Exercise and lung health. <https://www.lung.org/lung-health-diseases/wellness/exercise-and-lung-health>. [2024-07-11]
45. Soriano JB, Polverino F, Cosio BG. What is early COPD and why is it important? *Eur Respir J*, 2018; 52, 1801448.
46. Martinez FJ, Han MK, Allinson JP, et al. At the root: defining and halting progression of early chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*, 2018; 197, 1540–51.
47. Kohansal R, Martinez-Camblor P, Agustí A, et al. The natural history of chronic airflow obstruction revisited: an analysis of the Framingham offspring cohort. *Am J Respir Crit Care Med*, 2009; 180, 3–10.
48. Sanchez-Salcedo P, Divo M, Casanova C, et al. Disease progression in young patients with COPD: rethinking the Fletcher and Peto model. *Eur Respir J*, 2014; 44, 324–31.
49. Adeloye D, Song PG, Zhu YJ, et al. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. *Lancet Respir Med*, 2022; 10, 447–58.
50. Wang C, Xu JY, Yang L, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China pulmonary health [CPH] study): a national cross-sectional study. *Lancet*, 2018; 391, 1706–17.
51. Maselli DJ, Bhatt SP, Anzueto A, et al. Clinical epidemiology of COPD: insights from 10 years of the COPDGene study. *Chest*, 2019; 156, 228–38.
52. Ji HW, Gulati M, Huang TY, et al. Sex differences in association of physical activity with all-cause and cardiovascular mortality. *J Am Coll Cardiol*, 2024; 83, 783–93.
53. Lee K, Lee H, Lee K, et al. Gender differences in pulmonary function, physical activity, and quality of life of patients with COPD based on data from the Korea National Health and Nutrition Examination Survey 2015 to 2019 from the Perspective of Pulmonary Rehabilitation. *Medicine*, 2022; 101, e31413.