Supplementary Methods

Genotyping

Genotyping was done with the UK Biobank Axiom Array (Axiom_UKB_WCSG) by Affymetrix Axiom Genotyping Services (Affymetrix, Inc., Santa Clara, CA) from DNA extracted from mostly blood but in a few instances, saliva. As previously described[1], variants with missing rate>5% or Hardy-Weinberg p-value<1×10^{-6} (if MAF>5%) were excluded and participants with missing call rate>5%, identity-by-state distance>0.9, sex discrepancies, or of non-European ancestry were excluded. We excluded variants with MAF<5%. Imputation was performed using the 1000 Genomes Integrated Phase 1 version 3 reference panel (released March 2012) in IMPUTE2[2]. Variants with imputation quality score<0.8 were excluded. Ancestral genetic principal components were calculated using Eigenstrat[3].

Genetic risk scores

PLINK uses a p-value threshold to select the SNPs from the given summary results from the base dataset (in our case the meta-analysis results from UK Biobank and SpiroMeta consortium[4]) and clumps the SNPs based on distance and LD. Clumping identifies and selects the most significant SNP (i.e., lowest p-value) in each LD block which reduces the correlation between the remaining SNPs, while retaining SNPs with the strongest statistical evidence. At the p-value threshold of 5×10^{-9}, there were 29,202 SNPs, 23,728 SNPs, and 37,301 SNPs associated with FEV_1, FVC, and FEV_1/FVC, respectively, in the meta-analysis of UK Biobank and SpiroMeta consortium. Among these, 23,673 SNPs for FEV_1, 19,864 SNPs for FVC, and 29,704 SNPs for FEV_1/FVC were available in ALHS. After LD clumping, the remaining SNPs were coded as risk alleles based on the directions of their effect estimates in the UK Biobank-SpiroMeta meta-analysis results for each trait[4]. The weighted genetic risk scores were calculated as the weighted sum of the number of the risk alleles, using the effect estimates from the UK Biobank-SpiroMeta meta-analysis as the weights[4]. The weighted sum varies across subjects and is often much lower than the total number of SNPs involved because the weights are usually small in magnitude. Technically, a unit change in a GRS represents a change of 1 allele in the weighted sum. The units are not particularly important here though: the GRS could be rescaled (multiplied by an arbitrary constant), and the statistical significance of the GRS effects and GRS by exposure interactions would remain unchanged.

References

1. Wyss AB, Sofer T, Lee MK, et al. Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function. *Nat Commun* 2018;9(1):2976. doi: 10.1038/s41467-018-05369-0 [published Online First: 2018/08/01]

2. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS genetics* 2009;5(6):e1000529. doi: 10.1371/journal.pgen.1000529 [published Online First: 2009/06/23]
3. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38(8):904-9. doi: 10.1038/ng1847 [published Online First: 2006/07/25]
4. 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(3):481-93. doi: 10.1038/s41588-018-0321-7 [published Online First: 2019/02/26]

Supplementary Tables

Table E1: Interaction between smoking and genetic risk scores (GRSs) in relation to FEV_1 and FVC

Trait	Smoking status	N	Intercept[a]	Smoking Effect[b]	GRS Effect[c]	GRS X Smoking Interaction: difference in the effect of GRS per smoking category[d]	$P_{interaction}$[e]
FEV_1	Never	1884	2.511	-	-0.031	-	-
	Former	839	2.407	-0.104	-0.035	-0.004	0.709
	Current	121	2.097	-0.414	-0.031	-0.0002	0.995
FVC	Never	1884	3.258	-	-0.078	-	-
	Former	839	3.235	-0.023	-0.100	-0.022	0.357
	Current	121	2.978	-0.280	-0.069	0.009	0.860

[a]The intercept at each smoking category is the outcome (FEV_1 or FVC) value for a subject in that smoking category calculated at the mean value for all continuous variables in the model (GRS, age, age^2, height, height^2, weight (for FVC only) and 10 principal components) and at reference category for all categorical covariates (i.e. non-asthmatics, females, and residing at Iowa).

[b]The effect of smoking is obtained by subtracting the intercept value for never smoking from the intercept value for the smoking category in question, calculated at mean value for all continuous variables (GRS, age, age^2, height, height^2, weight (for FVC only) and 10 principal components) and at reference category for all categorical covariates (i.e. non-asthmatics, females, and residing at Iowa).

[c]The effect for the GRS is the individual slope for GRS for each exposure category which is interpretable as the difference in the outcome (FEV_1 or FVC) per unit increase in the GRS.

[d]The interaction effect between the GRS and smoking is the difference in the effect estimate for the GRS by smoking category which is calculated as the difference in the slope for the GRS for that smoking category relative to never smokers.

[e]The p-value for interaction between the GRS and each smoking category.

Table E2: Interaction between asthma and genetic risk scores (GRSs) in relation to FEV_1 and FVC

Trait	Asthma status	N	Intercept[a]	Asthma Effect[b]	GRS Effect[c]	GRS X Asthma Interaction: difference in the effect of GRS per asthma category[d]	$P_{interaction}$[e]
FEV_1	No	1803	2.456	-	-0.028	-	-
	Yes	1041	2.160	-0.296	-0.032	-0.004	0.761
FVC	No	1803	3.227	-	-0.090	-	-
	Yes	1041	3.071	-0.156	-0.069	0.021	0.315

[a]The intercept at each asthma category is the outcome (FEV_1 or FVC) value for a subject in that asthma category calculated at the mean value for all continuous variables in the model (GRS, age, age^2, height, height^2, weight (for FVC only), packyears, and 10 principal components) and at reference category for all categorical covariates (i.e. never smokers, females, and residing at Iowa).

[b]The effect of asthma is obtained by subtracting the intercept value for non-asthmatics from the intercept value for the asthmatics, calculated at the mean value for all continuous variables (GRS, age, age^2, height, height^2, weight (for FVC only), packyears, and 10 principal components) and at reference category for all categorical covariates (i.e. never smokers, females, and residing at Iowa).

[c]The effect for the GRS is the individual slope for GRS for each exposure category which is interpretable as the difference in the outcome (FEV_1 or FVC) per unit increase in the GRS.

[d]The interaction effect between the GRS and asthma is the difference in the effect estimate for the GRS by asthma category which is calculated as the difference in the slope for the GRS for asthmatics relative to non-asthmatics.

[e]The p-value for interaction between the GRS and asthma.

Table E3: Interaction between genetic risk score (GRS) and smoking status by asthma status in relation to FEV_1/FVC.

		FEV_1/FVC				
Exposure	**N**	**Intercept[a]**	**Smoking Effect[b]**	**GRS Effect[c]**	**GRS X Smoking Interaction: difference in the effect of GRS per smoking category[d]**	**$P_{3\text{-way interaction}}$[e]**
Asthmatics	1041					
Smoking						
Never	717	0.715	-	-0.004	-	
Former	294	0.694	-0.021	-0.008	-0.004	
Current	30	0.623	-0.092	-0.024	-0.020	
						Former: 0.798 Current: 0.037
Non-asthmatics	1803					
Smoking						
Never	1167	0.760	-	-0.003	-	
Former	545	0.738	-0.022	-0.005	-0.002	
Current	91	0.676	-0.084	-0.006	-0.003	

[a]For each asthma category, the intercept at a smoking category is the FEV_1/FVC value for a subject in that smoking category calculated at the mean value for all continuous variables in the model (GRS, age, age^2, height, height^2, and 10 principal components) and at reference category for all categorical covariates (i.e. females and residing at Iowa).

[b]For each asthma category, the effect of smoking is obtained by subtracting the intercept value for never smoking from the intercept value for the smoking category in question, calculated at mean value for all continuous variables (GRS, age, age^2, height, height^2, and 10 principal components) and at reference category for all categorical covariates (i.e. females and residing at Iowa).

[c]For each asthma category, the effect for the GRS is the individual slope for GRS for each exposure category which is interpretable as the difference in FEV_1/FVC per unit increase in the GRS.

[d]For each asthma category, the interaction effect between the GRS and smoking is the difference in the effect estimate for the GRS by smoking category which is calculated as the difference in the slope for the GRS for that smoking category relative to never smokers.

[e]The p-value for interaction between the GRS, asthma status and smoking status.

Table E4: Interaction between genetic risk score (GRS) and smoking status by gender in relation to FEV_1/FVC.

		FEV_1/FVC				
Exposure	**N**	**Intercept[a]**	**Smoking Effect[b]**	**GRS Effect[c]**	**GRS X Smoking Interaction: difference in the effect of GRS per smoking category[d]**	**$P_{3\text{-way interaction}}$[e]**
Females	1398					
Smoking						
Never	1052	0.757	-	-0.003	-	
Former	296	0.747	-0.010	-0.006	-0.003	
Current	50	0.676	-0.081	-0.019	-0.016	
						Former: 0.632 Current: 0.017
Males	1446					
Smoking						
Never	832	0.752	-	-0.004	-	
Former	543	0.722	-0.030	-0.006	-0.002	
Current	71	0.661	-0.091	-0.003	0.001	

[a]For each gender category, the intercept at a smoking category is the FEV_1/FVC value for a subject in that smoking category calculated at the mean value for all continuous variables in the model (GRS, age, age^2, height, height^2, and 10 principal components) and at reference category for all categorical covariates (i.e. non-asthmatics and residing at Iowa).

[b]For each gender category, the effect of smoking is obtained by subtracting the intercept value for never smoking from the intercept value for the smoking category in question, calculated at mean value for all continuous variables (GRS, age, age^2, height, height^2, and 10 principal components) and at reference category for all categorical covariates (i.e. non-asthmatics and residing at Iowa).

[c]For each gender category, the effect for the GRS is the individual slope for GRS for each exposure category which is interpretable as the difference in FEV_1/FVC per unit increase in the GRS.

[d]For each gender category, the interaction effect between the GRS and smoking is the difference in the effect estimate for the GRS by smoking category which is calculated as the difference in the slope for the GRS for that smoking category relative to never smokers.

[e]The p-value for interaction between the GRS, gender and smoking status.

Table E5: Interaction between log_{10}Endotoxin and genetic risk scores (GRSs) in relation to FEV_1 and FVC

Trait	**Endotoxin**	**N**	**Intercept[a]**	**log_{10}Endotoxin Effect[b]**	**GRS Effect[c]**	**GRS X log_{10}Endotoxin Interaction[d]**	**$P_{interaction}$[e]**
FEV_1	log_{10}Endotoxin	2208	2.478	-0.020	-0.033	0.001	0.932
FVC	log_{10}Endotoxin	2208	3.254	-0.005	-0.092	-0.0004	0.983

[a]The intercept is the outcome (FEV_1 or FVC) value for a subject calculated at the mean value for all continuous variables in the model (GRS, log_{10}Endotoxin, age, age^2, height, $height^2$, weight (for FVC only), packyears, and 10 principal components) and at reference category for all categorical covariates (i.e. never smokers, non-asthmatics, summer season of collection, females, and residing at Iowa).

[b]The effect of log_{10}Endotoxin is the difference in the outcome (FEV_1 or FVC) per unit increase in log_{10}Endotoxin, calculated at mean value for all continuous variables (GRS, age, age^2, height, $height^2$, weight (for FVC only), packyears, and 10 principal components) and at reference category for all categorical covariates (i.e. never smokers, non-asthmatics, summer season of collection, females, and residing at Iowa).

[c]The effect for the GRS is the slope for GRS which is interpretable as the difference in the outcome (FEV_1 or FVC) per unit increase in the GRS, calculated at mean value for all continuous variables (log_{10}Endotoxin, age, age^2, height, $height^2$, weight (for FVC only), packyears, and 10 principal components) and at reference category for all categorical covariates (i.e. never smokers, non-asthmatics, summer season of collection, females, and residing at Iowa).

[d]The interaction effect between the GRS and log_{10}Endotoxin is the difference in the effect estimate for the GRS per unit increase in log_{10}Endotoxin.

[e]The p-value for interaction between the GRS and log_{10}Endotoxin.

Table E6: Interaction between genetic risk score (GRS) and log_{10}Endotoxin by asthma status in relation to FEV_1/FVC.

		FEV_1/FVC				
Exposure	**N**	**Intercept[a]**	**log_{10}Endotoxin Effect[b]**	**GRS Effect[c]**	**GRS X log_{10}Endotoxin Interaction[d]**	**$P_{3\text{-way interaction}}$[e]**
Asthmatics						
log_{10}Endotoxin	817	0.707	-0.012	-0.007	-0.001	
						0.667
Non-asthmatics						
log_{10}Endotoxin	1391	0.754	-0.001	-0.003	-0.0005	

[a]For each asthma category, the intercept is the FEV_1/FVC value for a subject calculated at the mean value for all continuous variables in the model (GRS, log_{10}Endotoxin, age, age^2, height, height^2, packyears, and 10 principal components) and at reference category for all categorical covariates (i.e. never smokers, summer season of collection, females, and residing at Iowa).

[b]For each asthma category, the effect of log_{10}Endotoxin is interpretable as the difference in FEV_1/FVC per unit increase in the log_{10}Endotoxin calculated at mean value for all continuous variables (GRS, age, age^2, height, height^2, packyears, and 10 principal components) and at reference category for all categorical covariates (i.e. never smokers, summer season of collection, females, and residing at Iowa).

[c]For each asthma category, the effect for the GRS is the individual slope for GRS which is interpretable as the difference in FEV_1/FVC per unit increase in the GRS, calculated at the mean value for all continuous variables (log_{10}Endotoxin, age, age^2, height, height^2, packyears, and 10 principal components) and at reference category for all categorical covariates (i.e. never smokers, summer season of collection, females, and residing at Iowa).

[d]For each asthma category, the interaction effect between the GRS and log_{10}Endotoxin is the difference in the effect estimate for the GRS per unit increase in log_{10}Endotoxin.

[e]The p-value for interaction between the GRS, asthma status and log_{10}Endotoxin.

Supplementary Figures

Figure E1: Density plots of genetic risk scores (GRSs) for FEV_1, FVC, and FEV_1/FVC among 2844 study participants.

This figure shows the density plots of genetic risk scores (GRSs) for the three pulmonary function traits. The numbers of SNPs included in each GRS were 1123, 835, and 1691 for FEV_1, FVC, and FEV_1/FVC, respectively. The density values were obtained from kernel density estimation using the R function "density" with default settings, where a Gaussian kernel was used, and the smoothing parameter was chosen based on Silverman's rule of thumb.

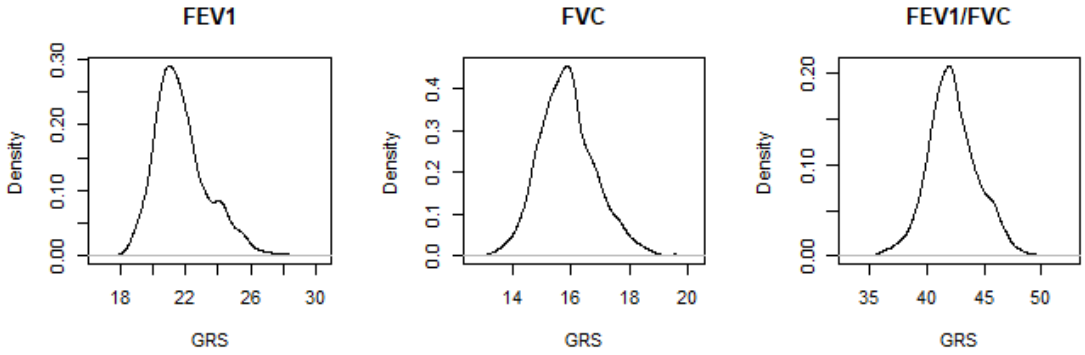

Figure E2: Association between GRS and FEV_1/FVC differs by smoking status and asthma status

This figure shows the estimated FEV_1/FVC values from the 3-way interaction model between smoking status, GRS, and asthma status, adjusting for age, age^2, height, height^2, state, gender, and 10 principal components. For each asthma category, the estimated FEV_1/FVC values are plotted against the range of GRS in our data for the three smoking categories, calculated at the mean values of all continuous covariates (age, age^2, height, height^2, and 10 principal components) and at reference category for all categorical covariates (i.e., females, and residing at Iowa). The shaded areas denote 95% pointwise confidence bands.

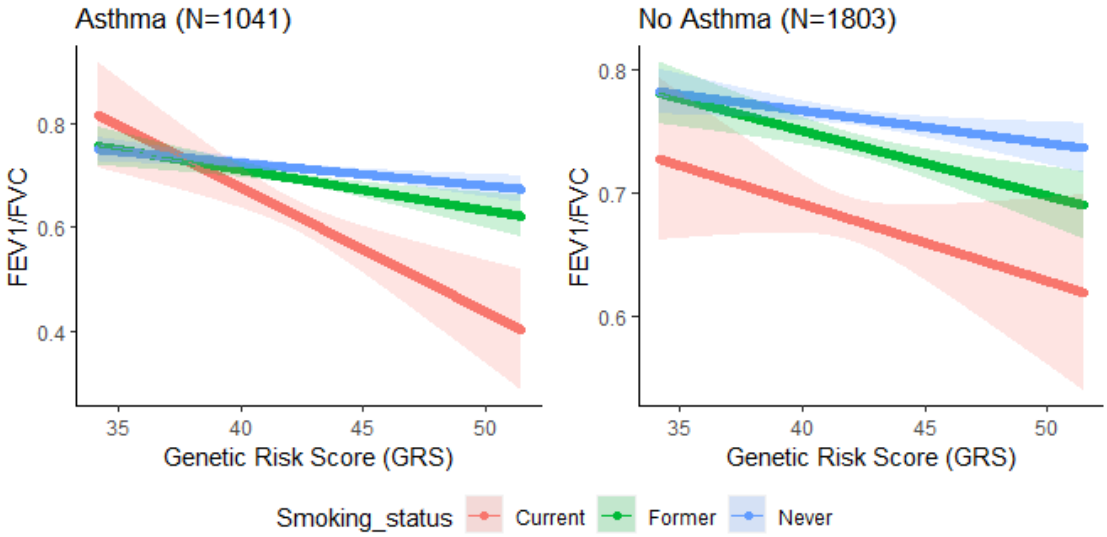

Figure E3: Association between GRS and FEV_1/FVC differs by smoking status and gender

This figure shows the estimated FEV_1/FVC values from the 3-way interaction model between smoking status, GRS, and gender, adjusting for age, age^2, height, height^2, state, asthma status, and 10 principal components. For each gender category, the estimated FEV_1/FVC values are plotted against the range of GRS in our data for the three smoking categories, calculated at the mean values of all continuous covariates (age, age^2, height, height^2, and 10 principal components) and at reference category for all categorical covariates (i.e., non-asthmatics, and residing at Iowa). The shaded areas denote 95% pointwise confidence bands.