2021

Interaction between Genetic Risk Scores for Reduced Pulmonary Function and Smoking, Asthma and Endotoxin

Sinjini Sikdar
Old Dominion University, ssikdar@odu.edu

Annah B. Wyss

Mi Kyeong Lee

Thanh T. Hoang

Marie Richards

See next page for additional authors

Original Publication Citation

Sikdar, S., Wyss, A.B., Lee, M.K., Hoang, T.T., Richards, M., Beane Freeman, L.E., Parks, C., Thorne, P.S., Hankinson, J.L., Umbach, D.M., Motsinger-Reif, A., & London, S.J. (2021). Interaction between Genetic Risk Scores for reduced pulmonary function and smoking, asthma and endotoxin. *Thorax 76*, 1219-1226. https://doi.org/10.1136/thoraxjnl-2020-215624

OPEN ACCESS

Original research

Interaction between Genetic Risk Scores for reduced pulmonary function and smoking, asthma and endotoxin

Sinjini Sikdar ,[1,2] Annah B Wyss,[2] Mi Kyeong Lee,[2] Thanh T Hoang,[2] Marie Richards,[3] Laura E Beane Freeman,[4] Christine Parks,[2] Peter S Thorne,[5] John L Hankinson,[6] David M Umbach,[7] Alison Motsinger-Reif,[7] Stephanie J London [2]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/thoraxjnl-2020-215624).

For numbered affiliations see end of article.

Correspondence to
Dr Stephanie J London, Epidemiology Branch, National Institute of Environmental Health Sciences, Durham, NC 27709, USA; london2@niehs.nih.gov

Received 23 June 2020
Revised 17 February 2021
Accepted 22 March 2021
Published Online First 7 May 2021

ABSTRACT

Rationale Genome-wide association studies (GWASs) have identified numerous loci associated with lower pulmonary function. Pulmonary function is strongly related to smoking and has also been associated with asthma and dust endotoxin. At the individual SNP level, genome-wide analyses of pulmonary function have not identified appreciable evidence for gene by environment interactions. Genetic Risk Scores (GRSs) may enhance power to identify gene–environment interactions, but studies are few.

Methods We analysed 2844 individuals of European ancestry with 1000 Genomes imputed GWAS data from a case–control study of adult asthma nested within a US agricultural cohort. Pulmonary function traits were FEV_1, FVC and FEV_1/FVC. Using data from a recent large meta-analysis of GWAS, we constructed a weighted GRS for each trait by combining the top (p value$<5\times10^{-9}$) genetic variants, after clumping based on distance (±250 kb) and linkage disequilibrium ($r^2=0.5$). We used linear regression, adjusting for relevant covariates, to estimate associations of each trait with its GRS and to assess interactions.

Results Each trait was highly significantly associated with its GRS (all three p values$<8.9\times10^{-8}$). The inverse association of the GRS with FEV_1/FVC was stronger for current smokers ($p_{interaction}=0.017$) or former smokers ($p_{interaction}=0.064$) when compared with never smokers and among asthmatics compared with non-asthmatics ($p_{interaction}=0.053$). No significant interactions were observed between any GRS and house dust endotoxin.

Conclusions Evaluation of interactions using GRSs supports a greater impact of increased genetic susceptibility on reduced pulmonary function in the presence of smoking or asthma.

Key messages

What is the key question?

► Whether the reduction in pulmonary function associated with increasing genetic susceptibility is enhanced or reduced by having exposures to smoking or house dust endotoxin or by having asthma.

What is the bottom line?

► Smoking or asthma amplifies the reduction in FEV_1/FVC that occurs with greater genetic susceptibility.

Why read on?

► Using the largest genome-wide association study meta-analysis of pulmonary function to date, we developed a robust Genetic Risk Score (GRS) for each pulmonary function trait in our data. We observed a significant interaction between the GRS for reduced FEV_1/FVC and smoking status. Our study is the first to examine interactions between GRSs for reduced pulmonary function and asthma status or house dust endotoxin exposure. We observed a marginally significant interaction between the GRS for reduced FEV_1/FVC and asthma. The finding that the association of genetic susceptibility with reduced pulmonary function is strongest among current smokers and asthmatics provides evidence that the population with higher genetic risk for impaired pulmonary function is more susceptible to the deleterious effects of smoking and asthma.

INTRODUCTION

Spirometric measures of pulmonary function, such as FEV_1, FVC and their ratio, FEV_1/FVC, are robust indices of respiratory health used in diagnosing and monitoring various lung conditions, including COPD. These pulmonary function metrics are predictors of mortality, even after adjusting for known risk factors.[1–4]

Pulmonary function is influenced by both genetic and environmental factors. Genome-wide association studies (GWASs) have identified many loci associated with pulmonary function.[5–9] Environmental exposures, most notably, cigarette smoking, also substantially influence pulmonary function.[10 11] Endotoxin, a lipopolysaccharide on the cell wall of Gram-negative bacteria ubiquitous in the environment, is a powerful initiator of innate immune response.[12] Occupational endotoxin exposure is associated with lower lung function.[13 14] Although endotoxin exposure in childhood might protect against asthma development,[15] in adulthood, endotoxin in house dust has been associated with lower pulmonary function in asthmatics.[16 17] Asthma is associated with reduced lung function in many studies.[18]

To cite: Sikdar S, Wyss AB, Lee MK, *et al*. *Thorax* 2021;**76**:1219–1226.

Considerable efforts to identify interactions between individual genetic variants and environmental exposures for many human traits and diseases have identified few to no significant interactions.[19–21] Even with large sample sizes, power is limited to detect interactions with individual single-nucleotide polymorphisms (SNPs) in genome-wide analyses.[19] Several authors have highlighted the advantage of using Genetic Risk Scores (GRSs) over individual SNPs for identifying significant interactions.[22–24] For example, a genome-wide meta-analysis of FEV_1 and FEV_1/FVC by Hancock *et al* of nearly 50 000 individuals incorporated interaction with smoking but identified no genome-wide significant interactions despite the well-established association of smoking with these phenotypes.[20] Using the summary results from Hancock *et al*[20] for 26 SNPs previously identified in main effects GWAS of pulmonary function,[7] Aschard *et al* performed single SNP-by-smoking interaction tests and found no significant interactions.[24] However, combining the effects of these individual SNPs into a GRS identified a significant interaction between smoking status and the GRS on FEV_1/FVC.[24] In a study of cotton textile workers, Zhang *et al* found a significant interaction between occupational endotoxin exposure and a 10-SNP GRS for lower FEV_1 for longitudinal decline in FEV_1.[25] We are not aware of studies examining whether associations of GRSs with pulmonary function differ by asthma status.

Recently, a large-scale meta-analysis involving around 400 000 participants of European ancestry from the UK Biobank and SpiroMeta consortium brought the number of loci for pulmonary function to nearly 300.[8] This largest meta-analysis of these outcomes to date, provides the ability to generate authoritative risk scores for pulmonary function in individuals of European ancestry. Shrine *et al* constructed a single GRS from variants identified for any of four pulmonary function traits (FEV_1, FVC, FEV_1/FVC and peak expiratory flow) weighted by the effect sizes for FEV_1/FVC but found no interaction of GRS with ever-never smoking in relation to FEV_1/FVC.[8]

We constructed GRSs for reduced pulmonary function based on results from the aforementioned meta-analysis[8] to investigate whether associated genetic risk for reduced pulmonary function is more pronounced in the presence of smoking or other exposures related to reduced pulmonary function. We constructed a separate GRS for each of the three spirometric traits (FEV_1, FVC and FEV_1/FVC) based on the meta-analysis[8] and applied these three GRSs in a case–control study of asthma in adults nested within a US farming cohort with data on smoking and house dust endotoxin. We examined an interaction hypothesis, namely, whether the reduction in pulmonary function associated with increasing GRS is enhanced or reduced by exposure to smoking or house dust endotoxin or by having asthma.

METHODS

Study population and pulmonary function

The Agricultural Lung Health Study (ALHS) is a case–control study of current asthma in farmers and spouses of farmers, nested within the Agricultural Health Study.[26] We enrolled 3301 participants in the ALHS from 2009 to 2013. Details regarding the ALHS study design, including measurement of pulmonary function, have been previously reported.[16 27 28] Briefly, pulmonary function (FEV_1 (in litres), FVC (in litres) and FEV_1/FVC (proportion)) was measured during home visits by trained field technicians in accordance with American Thoracic Society guidelines.[16 29] Tests were graded by Dr John Hankinson; participants with quality grades of D or F were excluded from analysis.[16 30]

Classification of asthma

As previously described,[16 27] asthma cases were identified from the larger Agricultural Health Study cohort in three categories: self-reported doctor-diagnosed current asthma, potential undiagnosed asthma based on the presence of current asthma symptoms and asthma medication use in non-smokers, and overlapping diagnoses of current asthma and either COPD or emphysema in non-smokers. A random sample of cohort members who did not meet any of these case definitions was selected for enrolment as non-cases.

Endotoxin measurements

House dust samples were collected by vacuuming bedroom floors and sleeping surfaces of participants.[16] Endotoxin levels in house dust were measured using the Limulus amebocyte lysate assay (Lonza Walkersville, Walkersville, Maryland, USA), as previously described.[31–33] Measurements below the limit of detection were assigned a value equal to that limit divided by the square root of two.

Assessment of smoking

Smoking history was obtained from questionnaires. Participants were classified as current, former or never smokers. Pack-years were calculated as packs smoked per day times years smoked.

Genotyping

Details about the genotyping, imputation and quality control are in the online supplemental material.

Genetic Risk Scores

Weighted GRSs were constructed using the complete summary results from the previous meta-analysis of more than 400 000 individuals of European ancestry.[8] The summary results were pruned for linkage disequilibrium (LD) using the p value informed clumping method in PLINK V.1.9,[34] based on the LD structure in the ALHS using a distance of ±250 kb and LD threshold of 0.5. We used a p value threshold of 5×10^{-9} to maximise stringency and for consistency with currently recommended genome-wide significance thresholds for resequencing analyses of individuals of European ancestry.[8 35] After LD clumping, the numbers of SNPs remaining for GRS calculation were 1123 for FEV_1, 835 for FVC and 1691 for FEV_1/FVC. Weighted GRSs for ALHS participants were calculated as the weighted sum of the number of the risk alleles using effect estimates from the UK Biobank-SpiroMeta meta-analysis as weights.[8] Further details about the calculation of GRSs can be found in the online supplemental material.

Statistical analyses

Using linear regression, we tested associations between each trait (FEV_1 (litres), FVC (litres) and FEV_1/FVC (proportion)) and its corresponding GRS adjusting for age, age^2, height, height^2, asthma (case and non-case), smoking status (current, former or never), pack-years of smoking, state of residence (Iowa and North Carolina), gender, first 10 genetic principal components and weight (kg, FVC only). Model examining associations of traits with smoking (two dummy variables for former or current smoking vs never) included the aforementioned covariates without pack-years or principal components. Models for association of traits with asthma were additionally adjusted for smoking and pack-years. Endotoxin was $\log_{10}$-transformed and models for association with traits were further adjusted for season of collection of dust sample. Interactions between the

GRS and each exposure (smoking, asthma or endotoxin) were tested by adding product terms to the aforementioned models and adjusting for the first 10 genetic principal components. Where we identified significant two-way interactions, we considered further three-way interaction terms with the remaining two exposures. We considered a nominal p value cut-off of 0.05 for statistical significance of our results. All analyses were performed in R.[36] Analyses used data release AHSREL201304.00.

RESULTS

Study participants

Among the 3301 ALHS participants, 3069 had spirometry passing quality control and complete data on smoking, asthma and covariates, including 2844 of European ancestry based on principal components analysis. Among these 2844 participants, 1041 were asthma cases. Current smoking was reported by 4.3% and former smoking by 29.5% (table 1). About 52% were farmers; the rest were spouses of farmers. House dust endotoxin measurements were available for 2385 participants. Among these, 177 visits were to homes where a spouse had already been enrolled; spouses were removed, leaving 2208 participants for analyses of endotoxin.

Table 1 Characteristics of the 2844 participants

Characteristics	n (%)
Gender	
Female	1398 (49.2)
Male	1446 (50.8)
Enrolment status	
Farmer	1491 (52.4)
Spouse	1353 (47.6)
State	
Iowa	2055 (72.3)
North Carolina	789 (27.7)
Current asthma status	
Case	1041 (36.6)
Non-case	1803 (63.4)
Smoking status	
Never	1884 (66.2)
Former	839 (29.5)
Current	121 (4.3)
Season of endotoxin measurement (n=2208)*	
Summer	628 (28.4)
Spring	586 (26.5)
Fall	492 (22.3)
Winter	502 (22.7)
FEV_1	
Median (25th–75th percentiles) (L)	2.5 (2.0–3.1)
FVC	
Median (25th–75th percentiles) (L)	3.4 (2.8–4.2)
FEV_1/FVC	
Median (25th–75th percentiles), proportion	0.75 (0.69–0.79)
Age	
Median (25th–75th percentiles) (years)	62.8 (54.8–71.3)
Pack-years in ever smokers	
Median (25th–75th percentiles)	9 (1.5–26.9)
Number of cigarettes per day in current smokers	
Median (25th–75th percentiles)	10 (5–20)
Endotoxin in house dust (n=2208)	
Median (25th–75th percentiles) (EU/mg)	43.5 (20.1–73.5)

*House dust endotoxin data were available for 2208 participants after removing the 177 for whom a visit was also made to a spouse.

Association between exposures and pulmonary function

As expected, smoking status was highly significantly associated with lower FEV_1 and FEV_1/FVC, with larger effect estimates for current smoking than for former smoking relative to never smoking (table 2). For FVC, inverse associations were observed for both current and former smoking, though the association for former smoking did not reach statistical significance (table 2). Pack-years of smoking was inversely associated with all three pulmonary function traits: FEV_1: β=—0.009 L/pack-year, p value$<2.0\times10^{-16}$; FEV_1/FVC: β=—0.002 L/pack-year, p value$<2.0\times10^{-16}$; and FVC: β=—0.005 L/pack-year, p value$=4.4\times10^{-12}$.

Asthma was highly statistically significantly associated with lower pulmonary function for all three traits (table 2). Log_{10}-transformed house dust endotoxin was inversely related to all three traits but not statistically significantly (table 2).

Genetic Risk Scores

Summary statistics of the GRSs for the three pulmonary function traits are shown in table 3; distributions of the GRSs are shown in online supplemental figure E1. As expected, the GRSs were highly significantly associated with lower values for each pulmonary function trait (table 3, all p values$<8.9\times10^{-8}$).

Interaction between GRSs and smoking status

We observed significant interactions between the GRS for FEV_1/FVC and smoking status. The interaction effect between GRS and smoking status shows the difference in the effects of GRS on FEV_1/FVC between smokers (current or former) and never smokers. The inverse association between GRS and FEV_1/FVC was greater for current smokers than never smokers (table 4); the estimated effect of GRS, per unit increase, on FEV_1/FVC in never smokers was −0.003 and that for the current smokers was −0.012 with a difference of −0.009 ($p_{interaction}=0.017$). Even for former smokers, the inverse association between GRS and FEV_1/FVC was higher compared with never smokers, where the estimated effect of GRS on FEV_1/FVC in former smokers was −0.006 vs −0.003 in never smokers, with a difference of −0.003 ($p_{nteraction}=0.064$). Figure 1 plots the association between the GRS and FEV_1/FVC according to smoking status and shows that the harmful effects of smoking were larger among participants with higher GRSs. No significant interactions with the GRS were seen with smoking for FEV_1 or FVC (lowest $p_{interaction}=0.357$, online supplemental table E1). We also tested for interactions between the GRS and pack-years of smoking in relation to each of the three traits but none were close to statistically significant (FEV_1: $\beta_{interaction}=0.0003$, $p_{interaction}=0.406$; FVC: $\beta_{interaction}=0.0005$, $p_{interaction}=0.480$; FEV_1/FVC: $\beta_{interaction}=-0.00002$, $p_{interaction}=0.674$). However, among current smokers, we observed a significant interaction between the GRS and the number of cigarettes smoked per day for FEV_1/FVC ($\beta_{interaction}=-0.0004$, $p_{interaction}=0.027$).

Given that there is some genetic contribution to smoking behaviour and we identified an interaction between the GRS

Table 2 Association between the exposures and pulmonary function traits

		FEV_1 (L)		FVC (L)		FEV_1/FVC	
Exposures	**n**	**β (SE)**	**P value**	**β (SE)**	**P value**	**β (SE)**	**P value**
Smoking							
Never	1884	Referent	–	Referent	–	Referent	–
Former	839	−0.108* (0.021)	4.2×10^{-7}	−0.028* (0.024)	0.246	−0.023* (0.003)	4.2×10^{-11}
Current	121	−0.412* (0.047)	$<2.0\times10^{-16}$	−0.280* (0.052)	9.1×10^{-8}	−0.084* (0.008)	$<2.0\times10^{-16}$
Asthma							
No	1803	Referent	–	Referent	–	Referent	–
Yes	1041	−0.297† (0.019)	$<2.0\times10^{-16}$	−0.155† (0.022)	1.1×10^{-12}	−0.048† (0.003)	$<2.0\times10^{-16}$
Endotoxin in house dust							
log_{10}endotoxin	2208	−0.017‡ (0.015)	0.254	−0.002‡ (0.017)	0.883	−0.004‡ (0.002)	0.131

*Estimates adjusted for age, age^2, state, gender, height, height^2 and asthma status (body weight for FVC only).
†Estimates adjusted for age, age^2, state, gender, height, height^2, smoking status and pack-years (body weight for FVC only).
‡Estimates adjusted for age, age^2, state, gender, height, height^2, asthma status, season of dust collection, smoking status, and pack-years (body weight for FVC only).

and smoking status in relation to FEV_1/FVC, we tested whether its GRS was related to smoking and found no appreciable association (adjusting for age, age^2, height, height^2, asthma status, state, gender and genetic principal components): former smokers β=0.162, SE=0.096, p value=0.090; current smokers β=−0.139, SE=0.210, p value=0.508.

Interaction between GRSs and asthma

We observed a marginally significant interaction between the GRS and asthma in relation to FEV_1/FVC with a stronger inverse association between the GRS and FEV_1/FVC in asthmatics (estimated effect of a unit increase in GRS on FEV_1/FVC=−0.006) than non-asthmatics (−0.003) with a $p_{interaction}$=0.053 (table 5). From figure 2, asthma had a stronger negative effect on FEV_1/FVC among participants with higher GRSs. No appreciable interaction with asthma was seen for FEV_1 or FVC (online supplemental table E2).

Given the interaction between asthma and the GRS in relation to FEV_1/FVC, we evaluated the association between asthma and the GRS for this trait, adjusting for age, age^2, height, height^2, smoking status, pack-years, state, gender and genetic principal components. The GRS for FEV_1/FVC was not significantly related to asthma (β=0.136, SE=0.087, p value=0.116).

Three-way interaction between smoking, asthma and GRSs

For FEV_1/FVC, we examined whether the interaction effect between GRS and smoking status differed by asthma. In asthmatics, FEV_1/FVC had steeper inverse relationship with increased genetic risk in current smokers (when compared with never smokers) than in non-asthmatics, yielding a statistically significant three-way interaction (online supplemental table E3 and online supplemental figure E2). The interaction effect between GRS and former smoking (in comparison to never smokers) for FEV_1/FVC was not significantly different between asthmatics and non-asthmatics (online supplemental table E3).

Three-way interaction between smoking, gender and GRSs

Additionally, we examined whether the interaction effect between GRS for FEV_1/FVC and smoking status differed by gender. In women, FEV_1/FVC had a steeper inverse relationship with increasing genetic risk in current smokers compared with never smokers, whereas no such difference between current and never smokers was observed in men, yielding a significant three-way interaction effect between GRS, gender and current smoking (vs never smoking) (online supplemental table E4 and figure E3). The interaction effect between GRS and former smoking (in comparison to never smokers) was not significantly different by gender (online supplemental table E4).

Interaction between GRSs and endotoxin

We observed no significant interactions between the GRS and endotoxin for any of the traits (table 6 and online supplemental table E5). Because we had previously reported a stronger association between endotoxin and FEV_1/FVC in asthmatics than non-asthmatics,[16] we evaluated a possible three-way interaction with asthma but found no evidence for one ($p_{three\text{-}way\ interaction}$=0.667, online supplemental table E6).

DISCUSSION

As expected, all three pulmonary function traits (FEV_1, FVC and FEV_1/FVC) were significantly lower among both current and former smokers compared with never smokers, and asthmatics had lower pulmonary function than non-asthmatics. We developed a separate GRS for each of the three pulmonary function traits in our study population using a large-scale meta-analysis

Table 3 Association between GRSs and pulmonary function traits

	GRS					
Outcome	**Range**	**Median (25th–75th percentiles)**	**Mean**	**GRS effect estimate***	**SE**	**P value**
FEV_1 (L)	17.5–29.5	21.5 (20.6–22.6)	21.8	−0.029	0.005	8.8×10^{-8}
FVC (L)	13.1–19.8	15.8 (15.2–16.5)	15.9	−0.082	0.010	5.7×10^{-15}
FEV_1/FVC	34.2–51.5	42.1 (40.9–43.6)	42.3	−0.004	0.001	1.0×10^{-9}

*Effect estimates provide the change in the trait (in litres for FEV_1 and FVC, proportion with range 0–1 for FEV_1/FVC) per one unit increase in the GRSs. Pulmonary function traits were regressed on the GRS for that trait, with adjustment for age, age^2, state, gender, height, height^2, asthma status, smoking status, pack-years, first 10 principal components, and for FVC only, body weight.
GRS, Genetic Risk Score.

Table 4 Interaction between smoking and GRS in relation to FEV_1/FVC

		FEV_1/FVC				
Exposure	**n**	**Intercept***	**Smoking effect†**	**GRS effect‡**	**GRS×smoking interaction: difference in the effect of GRS per smoking category§**	**$P_{interaction}$¶**
Smoking						
Never	1884	0.760	–	−0.003	–	–
Former	839	0.738	−0.022	−0.006	−0.003	0.064
Current	121	0.673	−0.087	−0.012	−0.009	0.017

*The intercept at each 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 the reference category for all categorical covariates (ie, non-asthmatic, female and residing at Iowa).
†The effect of smoking is obtained by subtracting the intercept value for never smoking from the intercept value for the smoking category in question. For example, for former smokers, 0.738–0.760=−0.022 is the difference in FEV_1/FVC for a former smoker relative to a never smoker calculated at the mean value for all continuous variables (GRS, age, age^2, height, height^2 and 10 principal components) and at the reference category for all categorical covariates (ie, non-asthmatic, female and residing at Iowa).
‡The effect for the GRS is the individual slope for that GRS for each exposure category and is interpretable as the difference in FEV_1/FVC per unit increase in the GRS.
§The interaction effect between the GRS and smoking is the difference in the effect estimate for that GRS by smoking category and is calculated as the difference in the slope for the GRS for that smoking category relative to never smokers. For former smokers this difference is −0.006−(−0.003)=−0.003.
¶The p value for interaction between the GRS and each smoking category.
GRS, Genetic Risk Score.

of European ancestry populations.[8] These GRSs were highly statistically significantly associated with lower values for their corresponding pulmonary function traits. We observed a significant interaction effect where the reduction in FEV_1/FVC with increasing GRS was more pronounced among current smokers and former than never smokers. We also found some evidence of interaction where the reduction in FEV_1/FVC with increasing GRS was more pronounced among asthmatics than among non-asthmatics.

Although statistical power is reduced for higher level interactions, we evaluated possible three-way interactions in situations where we identified significant two-way interactions. We found some evidence that the interaction between GRS and current smoking on reduced FEV_1/FVC was stronger among asthmatics

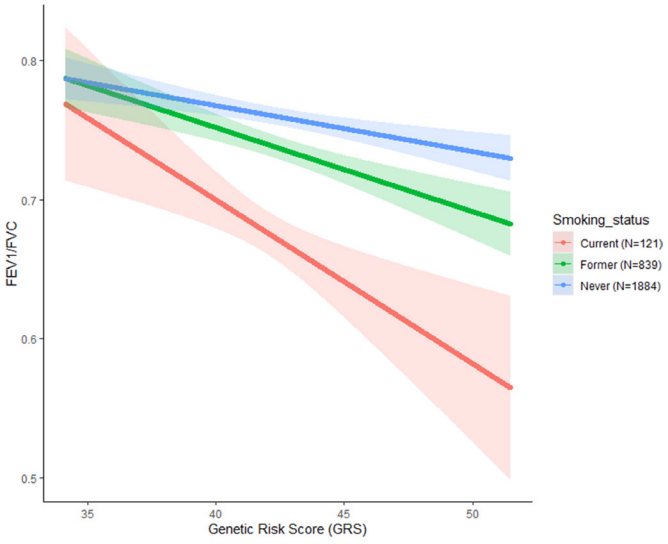

Figure 1 Association between GRS and FEV_1/FVC differs by smoking status. FEV_1/FVC is regressed on smoking status, GRS and their interaction, adjusting for age, age^2, height, height^2, state, gender, asthma status and 10 principal components. Shown are the estimated FEV_1/FVC values from the model against the range of GRS in our data for the three smoking categories (never, former and current), calculated at the mean values of all continuous variables (GRS, age, age^2, height, height^2 and 10 principal components) and at the reference category for all categorical covariates (ie, non-asthmatic, female and residing at Iowa). The shaded areas denote 95% pointwise confidence bands. GRS, Genetic Risk Score.

than non-asthmatics and among women than men. However, because of small numbers within these three-way cross-classified strata, interpretation of any significant three-way interactions requires caution.

For FEV_1/FVC, we observed significant interaction between its GRS and smoking status; for FEV_1 and FVC, we did not find interactions between their GRSs and smoking status. Results were similar for interactions between the GRSs and asthma status: present only for FEV_1/FVC. FEV_1/FVC is an index of airflow obstruction which is a characteristic of asthma and COPD and occurs with smoking.[37] Significant interactions between GRS and smoking or asthma for only FEV_1/FVC may reflect the fact that this parameter is independent of lung size. Genetic effects on FEV_1 and FVC, which reflect lung size, may have a predominant impact through lung development, which takes place largely in early life, rather than later response to environmental exposures or diseases. We also note that Aschard *et al*,[24] who examined both FEV_1 and FEV_1/FVC, identified an interaction between GRS and smoking predominantly for FEV_1/FVC.

To our knowledge, our study is the first to examine interactions between GRS of pulmonary function traits and asthma or house dust endotoxin exposure. Aschard *et al* found a significant interaction on FEV_1/FVC between an unweighted GRS based on 26 loci and ever versus never smoking, although this finding did not replicate in two independent datasets.[24] In the larger meta-analysis of Shrine *et al*, a single GRS based on 279 SNPs weighted by the effect sizes for FEV_1/FVC was constructed. That GRS did not interact with smoking status dichotomised as ever versus never.[8] We constructed a separate GRS based on the 278 of the 279 SNPs present in our data and tested for its interaction with smoking status (current, former vs never) in relation to FEV_1/FVC. The interaction effects with smoking status were not significant (former smokers: $p_{interaction}=0.11$, current smokers: $p_{interaction}=0.20$). However, using the more standard approach of creating a GRS based on clumping plus p value thresholding, we observed a significant interaction between our 1691-SNP GRS and smoking status in relation to FEV_1/FVC. This observation also highlights the advantage of using clumping plus p value thresholding to create a GRS over simple selection of top SNPs as discussed by Choi *et al*.[38] Shrine *et al* did not divide ever smokers into former and current for the interactions with GRS in their study. After several years from quitting, the decline in pulmonary function in former smokers tends to level off, so it is important to consider ever smokers in more detail. In our study, rather than creating just one weighted GRS, we created

Table 5 Interaction between asthma and GRS in relation to FEV_1/FVC

		FEV_1/FVC				
Exposure	n	Intercept*	Asthma effect†	GRS effect‡	GRS×asthma interaction: difference in the effect of GRS per asthma category§	$P_{interaction}$¶
Asthma						
No	1803	0.751	–	−0.003	–	–
Yes	1041	0.704	−0.047	−0.006	−0.003	0.053

*The intercept at each asthma category is the FEV_1/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, pack-years and 10 principal components) and at the reference category for all categorical covariates (ie, never smoker, female and residing at Iowa).
†The effect of asthma is obtained by subtracting the intercept value for non-asthmatics from the intercept value for the asthmatics; that is, 0.704–0.751=−0.047 is the difference in FEV_1/FVC for an asthmatic relative to a non-asthmatic calculated at the mean value for all continuous variables (GRS, age, age^2, height, height^2, pack-years and 10 principal components) and at the reference category for all categorical covariates (ie, never smoker, female and residing at Iowa).
‡The effect for the GRS is the individual slope for GRS for each exposure category and is interpretable as the difference in FEV_1/FVC per unit increase in the GRS.
§The interaction effect between the GRS and asthma is the difference in the effect estimate for the GRS by asthma category and is calculated as the difference in the slope for the GRS for asthmatics relative to non-asthmatics; that is, −0.006−(−0.003)=−0.003.
¶The p value for interaction between the GRS and asthma.
GRS, Genetic Risk Score.

a separate GRS for each pulmonary function trait weighted by the effect sizes for that trait. Our GRSs were based on the same large comprehensive GWAS meta-analysis as Shrine *et al*,[8] and we found evidence of interaction with smoking considering former and current smokers separately. The interaction was most notable in our data for current smokers relative to never smokers.

Our study has some limitations. Because asthma was categorised based on questionnaires, misclassification with COPD is possible. We did not adjust for socioeconomic status (SES). Occupation is often used to adjust for SES. Our participants were enrolled in the parent cohort because they were either farmers or spouses of farmers. By sharing an occupation, they would be regarded as having similar SES. Nevertheless, when we considered education as an alternate proxy for SES, the results did not materially change. Consistent with other genetic studies of pulmonary function, we did not adjust for comorbidities. However, if insufficient adjustment for SES or comorbidities can bias estimates of interaction with the GRS, we cannot exclude the possibility that this occurred. Because this is an agricultural population, participants potentially had higher exposure to endotoxin than the general US population. Additionally, all participants in this study and those in the UK Biobank and the SpiroMeta consortium were of European ancestry. Further, variants included in the GRS have different directions of associations with the pulmonary function traits. Although we recoded these directions to be uniform, combining the variants into a GRS might lose some information. However, assessing interactions using GRS provides greater statistical power than using individual variants.

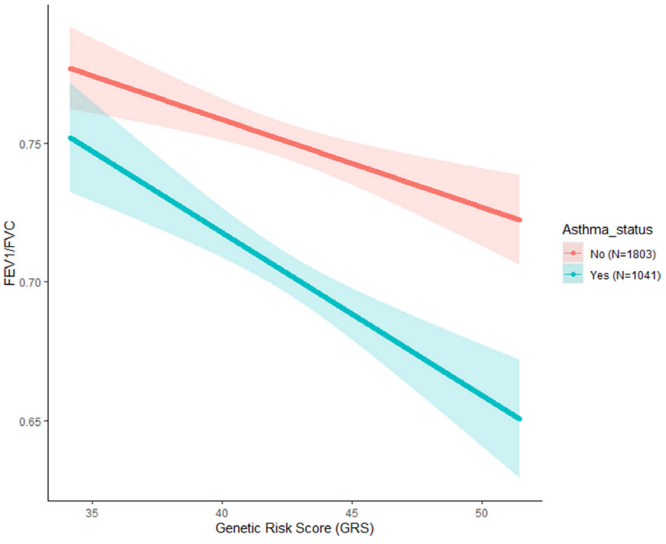

Figure 2 Association between GRS and FEV_1/FVC differs by asthma status. FEV_1/FVC is regressed on asthma status, GRS and their interaction, adjusting for age, age^2, height, height^2, state, gender, smoking status, pack-years, and 10 principal components. Shown are the estimated FEV_1/FVC values from the model against the range of GRS in our data for the two asthma categories, calculated at the mean values of all continuous variables (GRS, age, age^2, height, height^2, pack-years and 10 principal components) and 0 value for all categorical covariates (ie, never smoker, female and residing at Iowa). The shaded areas denote 95% pointwise confidence bands. GRS, Genetic Risk Score.

In most GWAS of pulmonary function, even though multiple correlated traits are examined simultaneously, correction for multiple testing based on the number of traits examined is not usually done.[8 9] There are few GWASs focusing on interaction hypotheses. We used a nominal p value of 0.05 for reporting significant interactions. If one were to adjust interaction p values for the three traits and three exposures considered, the p value threshold would be 0.05/9=0.006. At this stricter correction, none of our interaction findings would be significant. Thus, caution is required in the interpretation of our results pending replication in future studies.

A strength of the study is that we developed a separate GRS for each pulmonary function trait using a meta-analysis involving around 400 000 participants of European ancestry,[8] the largest GWAS of pulmonary function to date. This large-scale meta-analysis enabled generation of authoritative risk scores for pulmonary function in ALHS; we used these to investigate whether reduced pulmonary function associated with genetic risk is magnified in the presence of smoking or other exposures that have been related to reduced pulmonary function.

In conclusion, we developed separate GRSs for three pulmonary function traits in our study of asthma nested within an agricultural cohort. We identified significant interactions for FEV_1/FVC between its GRS and smoking status and marginally significant interactions for FEV_1/FVC between its GRS and asthma. Our data support the use of GRS to identify environmental interactions with genetic susceptibility. Although small numbers induced by further stratification require caution, we saw some evidence that, for FEV_1/FVC, the interaction between its GRS and smoking status differed by asthma and by gender. While it has been difficult to identify appreciable evidence of gene by environment interactions in genome-wide analyses at the individual

Table 6 Interaction between log_{10}endotoxin and GRS in relation to FEV_1/FVC

		FEV_1/FVC				
Exposure	**n**	**Intercept***	**log_{10}endotoxin effect†**	**GRS effect‡**	**GRS×log_{10}endotoxin interaction§**	**$P_{interaction}$¶**
Endotoxin						
log_{10}endotoxin	2208	0.754	−0.004	−0.004	−0.001	0.248

*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, $height^2$, pack-years and 10 principal components) and at the reference category for all categorical covariates (ie, never smoker, non-asthmatic, summer season of collection, female and residing at Iowa).
†The effect of log_{10}endotoxin is the difference in FEV_1/FVC per unit increase in log_{10}endotoxin, calculated at the mean value for all continuous variables (GRS, age, age^2, height, $height^2$, pack-years and 10 principal components) and at the reference category for all categorical covariates (ie, never smoker, non-asthmatic, summer season of collection, female and residing at Iowa).
‡The effect for the GRS is the slope for the 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, $height^2$, pack-years and 10 principal components) and at the reference category for all categorical covariates (ie, never smoker, non-asthmatic, summer season of collection, female and residing at Iowa).
§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.
¶The p value for interaction between the GRS and log_{10}endotoxin.
GRS, Genetic Risk Score.

SNP level, combining data across SNPs from large-scale GWAS through the use of GRSs can identify such interactions. Using the GRS approach, we find evidence that the impact of genetic susceptibility on reduced FEV_1/FVC is enhanced in the presence of smoking or asthma. These findings provide evidence that the population with higher genetic risk for impaired pulmonary function is more susceptible to the deleterious effects of smoking and asthma. Our findings might hint at potential biological mechanisms underlying the interactions between genetic variants and exposure to smoking, or presence of asthma, in relation to lung function. For example, significant interactions between genetic risk for reduced pulmonary function and smoking might suggest that some SNPs related to pulmonary function operate by influencing pathways for response to smoking, even though previous analyses of interaction with the individual SNPs have not identified significant interactions. Studies incorporating additional types of omics data, including proteomics and metabolomics, might help shed light on possible mechanisms. Future studies assessing interaction between GRSs and factors related to reduced pulmonary function would help to support stronger inferences regarding potential relevance in clinical practice.

Author affiliations
[1]Department of Mathematics and Statistics, Old Dominion University, Norfolk, Virginia, USA
[2]Epidemiology Branch, National Institute of Environmental Health Sciences, Durham, North Carolina, USA
[3]Westat, Durham, North Carolina, USA
[4]Occupational and Environmental Epidemiology Branch, National Cancer Institute, Bethesda, Maryland, USA
[5]Department of Occupational and Environmental Health, University of Iowa, Iowa City, Iowa, USA
[6]Hankinson Consulting, Athens, Georgia, USA
[7]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, Durham, North Carolina, USA

Acknowledgements We thank Dr Jane Hoppin, ScD (North Carolina State University, Raleigh North Carolina), for her important contribution to the Agricultural Lung Health Study during her tenure at the National Institute of Environmental Health Sciences (NIEHS); the numerous study staff at Social & Scientific Systems, Inc, who played a role in the data collection; Dr Frank Day of NIEHS and Dr Jianping Jin of Westat, Inc (Durham, North Carolina), for computational support; and Huiling Li of NIEHS for careful review of the manuscript.

Contributors SS and SJL designed the study and drafted the manuscript. SS performed data analysis with contributions from ABW, MKL, TTH and MR. JLH was responsible for quality control of the pulmonary function data. PST was responsible for the endotoxin measurements. DMU and AM-R provided statistical guidance. All authors provided critical input on the manuscript.

Funding This work was supported by the Intramural Research Programme of the National Institutes of Health, National Institute of Environmental Health Sciences (Z01-ES102385, Z01-ES049030, Z01-ES043012 and, for ABW, contract number HHSN273201600003I) and the National Cancer Institute (Z01-CP010119), and was supported in part by American Recovery and Reinvestment Act funds through National Institute of Environmental Health Sciences (contract number N01-ES-55546). PST was supported by PHR-SUPS2-S-10–00179 and NIH P30 ES005605.

Competing interests AM-R and PST report grants from the National Institute of Environmental Health Sciences during the conduct of the study.

Patient consent for publication Not required.

Ethics approval The study was approved by the institutional review board at the National Institute of Environmental Health Sciences.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request pending all required approvals.

ORCID iDs
Sinjini Sikdar http://orcid.org/0000-0003-1230-5162
Stephanie J London http://orcid.org/0000-0003-4911-5290

REFERENCES

1 Hole DJ, Watt GC, Davey-Smith G, *et al*. Impaired lung function and mortality risk in men and women: findings from the Renfrew and Paisley prospective population study. *BMJ* 1996;313:711–5.
2 Burney PGJ, Hooper R. Forced vital capacity, airway obstruction and survival in a general population sample from the USA. *Thorax* 2011;66:49–54.
3 Young RP, Hopkins R, Eaton TE. Forced expiratory volume in one second: not just a lung function test but a marker of premature death from all causes. *Eur Respir J* 2007;30:616–22.
4 Mannino DM, Buist AS, Petty TL, *et al*. Lung function and mortality in the United States: data from the first National health and nutrition examination survey follow up study. *Thorax* 2003;58:388–93.
5 Hancock DB, Eijgelsheim M, Wilk JB, *et al*. Meta-Analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet* 2010;42:45–52.
6 Loth DW, Soler Artigas M, Gharib SA, *et al*. Genome-Wide association analysis identifies six new loci associated with forced vital capacity. *Nat Genet* 2014;46:669–77.
7 Soler Artigas M, Loth DW, Wain LV, *et al*. Genome-Wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat Genet* 2011;43:1082–90.
8 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.
9 Wyss AB, Sofer T, Lee MK, *et al*. Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function. *Nat Commun* 2018;9:2976.
10 Urrutia I, Capelastegui A, Quintana JM, *et al*. Smoking habit, respiratory symptoms and lung function in young adults. *Eur J Public Health* 2005;15:160–5.
11 Burchfiel CM, Marcus EB, Curb JD, *et al*. Effects of smoking and smoking cessation on longitudinal decline in pulmonary function. *Am J Respir Crit Care Med* 1995;151:1778–85.
12 Nijland R, Hofland T, van Strijp JAG. Recognition of LPS by TLR4: potential for anti-inflammatory therapies. *Mar Drugs* 2014;12:4260–73.
13 Lai PS, Hang J-Q, Valeri L, *et al*. Endotoxin and gender modify lung function recovery after occupational organic dust exposure: a 30-year study. *Occup Environ Med* 2015;72:546–52.

14 Smid T, Heederik D, Houba R, *et al*. Dust- and endotoxin-related acute lung function changes and work-related symptoms in workers in the animal feed industry. *Am J Ind Med* 1994;25:877–88.
15 Braun-Fahrländer C, Riedler J, Herz U, *et al*. Environmental exposure to endotoxin and its relation to asthma in school-age children. *N Engl J Med* 2002;347:869–77.
16 Carnes MU, Hoppin JA, Metwali N, *et al*. House dust endotoxin levels are associated with adult asthma in a U.S. farming population. *Ann Am Thorac Soc* 2017;14:324–31.
17 Michel O, Kips J, Duchateau J, *et al*. Severity of asthma is related to endotoxin in house dust. *Am J Respir Crit Care Med* 1996;154:1641–6.
18 Sears MR. Lung function decline in asthma. *Eur Respir J* 2007;30:411–3.
19 Aschard H, Lutz S, Maus B, *et al*. Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. *Hum Genet* 2012;131:1591–613.
20 Hancock DB, Soler Artigas M, Gharib SA, *et al*. Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. *PLoS Genet* 2012;8:e1003098.
21 Ege MJ, Strachan DP, Cookson WOCM, *et al*. Gene-environment interaction for childhood asthma and exposure to farming in central Europe. *J Allergy Clin Immunol* 2011;127:138–44.
22 Ahmad S, Rukh G, Varga TV, *et al*. Gene × physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. *PLoS Genet* 2013;9:e1003607.
23 Qi Q, Chu AY, Kang JH, *et al*. Sugar-sweetened beverages and genetic risk of obesity. *N Engl J Med* 2012;367:1387–96.
24 Aschard H, Tobin MD, Hancock DB, *et al*. Evidence for large-scale gene-by-smoking interaction effects on pulmonary function. *Int J Epidemiol* 2017;46:dyw318–904.
25 Zhang R, Zhao Y, Chu M, *et al*. A large scale gene-centric association study of lung function in newly-hired female cotton textile workers with endotoxin exposure. *PLoS One* 2013;8:e59035.
26 Alavanja MC, Sandler DP, McMaster SB, *et al*. The agricultural health study. *Environ Health Perspect* 1996;104:362–9.
27 House JS, Wyss AB, Hoppin JA, *et al*. Early-life farm exposures and adult asthma and atopy in the agricultural lung health study. *J Allergy Clin Immunol* 2017;140:249–56.
28 Wyss AB, House JS, Hoppin JA, *et al*. Raw milk consumption and other early-life farm exposures and adult pulmonary function in the agricultural lung health study. *Thorax* 2018;73:279–82.
29 Miller MR, Hankinson J, Brusasco V, *et al*. Standardisation of spirometry. *Eur Respir J* 2005;26:319–38.
30 Hankinson JL, Eschenbacher B, Townsend M, *et al*. Use of forced vital capacity and forced expiratory volume in 1 second quality criteria for determining a valid test. *Eur Respir J* 2015;45:1283–92.
31 Vojta PJ, Friedman W, Marker DA, *et al*. First national survey of lead and allergens in housing: survey design and methods for the allergen and endotoxin components. *Environ Health Perspect* 2002;110:527–32.
32 Thorne PS, Kulhánková K, Yin M, *et al*. Endotoxin exposure is a risk factor for asthma: the National survey of endotoxin in United States housing. *Am J Respir Crit Care Med* 2005;172:1371–7.
33 Thorne PS, Mendy A, Metwali N, *et al*. Endotoxin exposure: predictors and prevalence of associated asthma outcomes in the United States. *Am J Respir Crit Care Med* 2015;192:1287–97.
34 Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–75.
35 Pulit SL, de With SAJ, de Bakker PIW. Resetting the bar: statistical significance in whole-genome sequencing-based association studies of global populations. *Genet Epidemiol* 2017;41:145–51.
36 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2019.
37 Contoli M, Baraldo S, Marku B, *et al*. Fixed airflow obstruction due to asthma or chronic obstructive pulmonary disease: 5-year follow-up. *J Allergy Clin Immunol* 2010;125:830–7.
38 Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc* 2020;15:2759–72.