Genome wide association study of Preserved Ratio Impaired Spirometry (PRISm)

Genome-wide association study of preserved ratio impaired spirometry (PRISm)

Daniel H. Higbee[1,2], Alvin Lirio[3], Fergus Hamilton[1], Raquel Granell[1], Annah B. Wyss[4], Stephanie J. London[4], Traci M. Bartz[5], Sina A. Gharib[6], Michael H. Cho[7,8], Emily Wan[7,8,9], Edwin Silverman[9], James D. Crapo[10], Jesus V.T. Lominchar[11], Torben Hansen[11], Niels Grarup[11], Thomas Dantoft[12], Line Kårhus[12], Allan Linneberg[12], George T. O'Connor[13,14], Josée Dupuis[15], Hanfie Xu[16], Maaike M. De Vries[17,18], Xiaowei Hu[19], Stephen S. Rich[19], R. Graham Barr[20], Ani Manichaikul[19], Sara R.A. Wijnant[21,22,23], Guy G. Brusselle[23,24], Lies Lahousse[21,22], Xuan Li[25], Ana I. Hernández Cordero[25], Ma'en Obeidat[25], Don D. Sin[25,26], Sarah E. Harris[27], Paul Redmond[27], Adele M. Taylor[27], Simon R. Cox[27], Alexander T. Williams[3], Nick Shrine[3], Catherine John[3], Anna L. Guyatt[3], Ian P. Hall[28], George Davey Smith[1], Martin D. Tobin[3,29,30] and James W. Dodd[1,2,30]

[1]MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, UK. [2]Academic Respiratory Unit, University of Bristol, Southmead Hospital, Bristol, UK. [3]Department of Population Health Sciences, University of Leicester, Leicester, UK. [4]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA. [5]Cardiovascular Health Research Unit, Departments of Biostatistics and Medicine, University of Washington, Seattle, WA, USA. [6]Computational Medicine Core, Center for Lung Biology, UW Medicine Sleep Center, Department of Medicine, University of Washington, Seattle, WA, USA. [7]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA. [8]Harvard Medical School, Boston, MA, USA. [9]Pulmonary and Critical Care Section, Department of Medicine, VA Boston Healthcare System, Boston, MA, USA. [10]Division of Pulmonary, Critical Care and Sleep Medicine, National Jewish Health, Denver, CO, USA. [11]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [12]Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Frederiksberg, Denmark. [13]Pulmonary Center, School of Medicine, Boston University, Boston, MA, USA. [14]Division of Pulmonary, Allergy, Sleep, and Critical Care Medicine, Boston Medical Center, Boston, MA, USA. [15]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada. [16]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. [17]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [18]Groningen Research Institute for Asthma and COPD, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [19]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA. [20]Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, NY, USA. [21]Department of Bioanalysis, Ghent University, Ghent, Belgium. [22]Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands. [23]Department of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium. [24]Department of Epidemiology, Department of Respiratory Medicine, Erasmus MC, Rotterdam, The Netherlands. [25]Centre for Heart Lung Innovation, University of British Columbia, St. Paul's Hospital, Vancouver, BC, Canada. [26]Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, BC, Canada. [27]Lothian Birth Cohorts, Department of Psychology, University of Edinburgh, Edinburgh, UK. [28]University of Nottingham and NIHR Nottingham Biomedical Research Centre, Nottingham, UK. [29]Leicester NIHR Biomedical Research Centre, Leicester, UK. [30]Joint senior authors.

Corresponding author: James W. Dodd (james.dodd@bristol.ac.uk)

Shareable abstract (@ERSpublications)
This is the first GWAS to report genome-wide significant SNPs for PRISm, four of which are novel for lung function. Genetic factors associated with PRISm are strongly correlated with risk of both other lung diseases and extrapulmonary comorbidity. https://bit.ly/3Qo0jUn

Cite this article as: Higbee DH, Lirio A, Hamilton F, *et al*. Genome-wide association study of preserved ratio impaired spirometry (PRISm). *Eur Respir J* 2024; 63: 2300337 [DOI: 10.1183/13993003.00337-2023].

Abstract

Background Preserved ratio impaired spirometry (PRISm) is defined as a forced expiratory volume in 1 s (FEV_1) <80% predicted and FEV_1/forced vital capacity ⩾0.70. PRISm is associated with respiratory symptoms and comorbidities. Our objective was to discover novel genetic signals for PRISm and see if they provide insight into the pathogenesis of PRISm and associated comorbidities.

Methods We undertook a genome-wide association study (GWAS) of PRISm in UK Biobank participants (Stage 1), and selected single nucleotide polymorphisms (SNPs) reaching genome-wide significance for

Received: 2 March 2023
Accepted: 29 Oct 2023

replication in 13 cohorts (Stage 2). A combined meta-analysis of Stage 1 and Stage 2 was done to determine top SNPs. We used cross-trait linkage disequilibrium score regression to estimate genome-wide genetic correlation between PRISm and pulmonary and extrapulmonary traits. Phenome-wide association studies of top SNPs were performed.

Results 22 signals reached significance in the joint meta-analysis, including four signals novel for lung function. A strong genome-wide genetic correlation (r_g) between PRISm and spirometric COPD (r_g=0.62, p<0.001) was observed, and genetic correlation with type 2 diabetes (r_g=0.12, p=0.007). Phenome-wide association studies showed that 18 of 22 signals were associated with diabetic traits and seven with blood pressure traits.

Conclusion This is the first GWAS to successfully identify SNPs associated with PRISm. Four of the signals, rs7652391 (nearest gene *MECOM*), rs9431040 (*HLX*), rs62018863 (*TMEM114*) and rs185937162 (*HLA-B*), have not been described in association with lung function before, demonstrating the utility of using different lung function phenotypes in GWAS. Genetic factors associated with PRISm are strongly correlated with risk of both other lung diseases and extrapulmonary comorbidity.

Introduction

Preserved ratio impaired spirometry (PRISm), also referred to as "restrictive pattern" or "unclassified" spirometry, is defined as forced expiratory volume in 1 s (FEV_1) <80% predicted and FEV_1/forced vital capacity (FVC) ratio ⩾0.70 [1]. It has been suggested that for a subgroup of subjects, PRISm may be a precursor of COPD, with up to 50% progressing to COPD while 15% return to "normal" spirometry over 5 years [2, 3]. A larger and younger cohort has shown that PRISm may be transient with only 12% going on to develop airflow obstruction over 8 years [4]. Clinical interest in PRISm relates to its consistent association with respiratory symptoms, comorbidities (*e.g.* obesity, diabetes and cardiovascular disease) and all-cause mortality [2–4].

Previous studies have shown that lung function measures or traits are, in part, heritable and associated with genetic variants, implicating a wide range of mechanisms including cilia development and elastic fibres in obstructive lung disease [5, 6], but the individual genetic associations and pathways which underlie PRISm are less well understood. A previous genome-wide association study (GWAS) of PRISm failed to find associations of genome-wide significance ($p<5\times10^{-8}$), was modest, and was restricted to ever-smokers [7].

Genetic variants associated with PRISm could provide invaluable insight into its pathogenesis and associated comorbidities, as well as potentially identify therapeutic targets.

Our objective was to perform a case–control GWAS of PRISm and report novel associated single nucleotide polymorphisms (SNPs) in a two-stage study design see if they provide insight into the pathogenesis of PRISm and associated comorbidities.

Methods

Study design

We performed a two-stage GWAS with meta-analysis. For the discovery cohort (Stage 1), we used the UK Biobank (UKBB) (www.ukbiobank.ac.uk). For Stage 2 we used cohorts within the SpiroMETA and Cohort for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortia as well as COPDGene. We estimated genetic correlations with potentially related phenotypic traits. We performed a phenome-wide association study (PheWAS) of SNPs not found to be associated with lung function in the largest lung function GWAS to date [8].

Stage 1

The UKBB, which was used for the Stage 1 GWAS, is a large UK population-based health research resource of ∼500 000 people aged 38–73 years old recruited between 2006 and 2010. Questionnaires, interviews, anthropometric measures and biological samples were collected. UKBB received ethical approval from the Research Ethics Committee (REC reference for UK Biobank: 11/NW/0382). We used previously derived variables of quality-controlled prebronchodilator FEV_1 and FVC. Only participants with spirometry classified as acceptable were included (supplementary appendix 1). Only those of self-identified European ancestry with very similar genetic ancestry based on principal component analysis of genotypes were included. Patients with unknown smoking status or weight were excluded. FEV_1 % predicted was calculated as per Global Lung Function Initiative (GLI) 2012 values using RSpiro R package in R v3.6.1 (www.r-project.org). PRISm was defined as FEV_1 <80% predicted and FEV_1/FVC ratio ⩾0.70 and controls as FEV_1 ⩾80% predicted and FEV_1/FVC ratio ⩾0.70. Participants with spirometry not meeting the criteria for PRISm or control were excluded. Figure 1 shows the participant selection flow chart and table 1 contains demographics of the sample used.

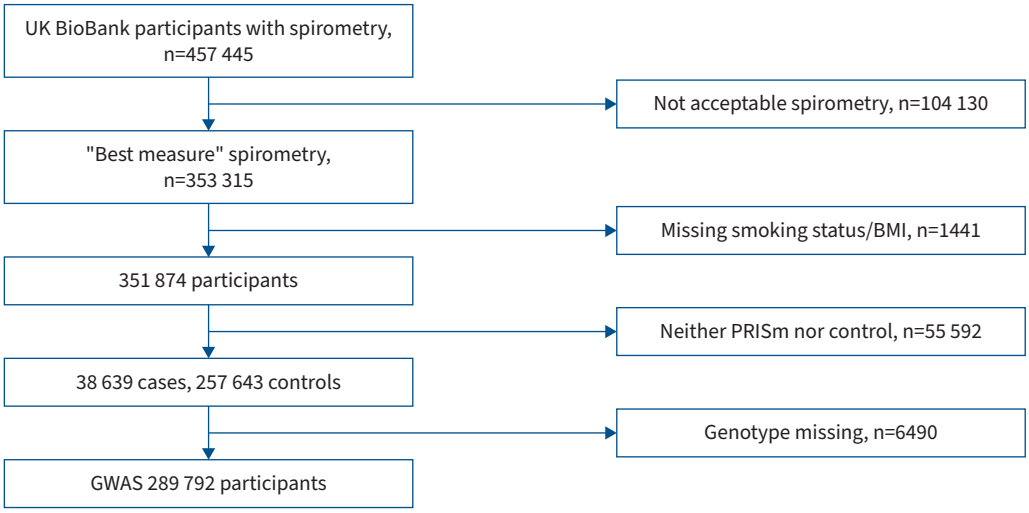

FIGURE 1 Participant selection flow chart. BMI: body mass index; PRISm: preserved ratio impaired spirometry; GWAS: genome-wide association study.

The GWAS was performed using the Integrative Epidemiology Unit (IEU) GWAS pipeline, described in detail elsewhere [9]. The pipeline contains previously derived genetic files of participants after pre-imputation quality control, phasing and imputation, allowing fast standardised GWAS of the UKBB population. A description of the process to create the derived genetic files can be found in detail elsewhere [10, 11] and can be summarised as follows. Genotyping was performed using the Axiom UK BiLEVE array and the Axiom Biobank array (Affymetrix) [12]. Before phasing, multiallelic SNPs or those with minor allele frequency (MAF) ⩽1% were removed. Phasing of genotype data was performed using a modified version of the SHAPEIT2 algorithm [13]. Genotype imputation to a reference set combining the UK10K haplotype and Haplotype Reference Consortium reference panels was performed using IMPUTE2 algorithms [14]. The analyses were restricted to autosomal variants using graded filtering with varying imputation quality for different allele frequency ranges. An in-house algorithm was then applied to preferentially remove the individuals related to the greatest number of other individuals until no related pairs remain. To model population structure in the sample, 143 006 directly genotyped SNPs were used, obtained after omitting variants with MAF <0.01, genotyping missing rate >0.015 or Hardy–Weinberg equilibrium p<0.0001; and linkage disequilibrium (LD) pruning to an r^2 threshold of 0.1 using PLINK v2.00 (www.cog-genomics.org/plink/2.0).

TABLE 1 Demographics of Stage 1 participants in UK Biobank

Demographic at baseline	PRISm	Controls
Participants (n)	38 639	257 643
Age (years)	56.4±7	56.0±7
BMI (kg·m^{-2})	29.1±5	27.2±4
Female (%)	55.4	55.6
FEV_1 (% predicted)	74 (68–77)	98 (90–106)
FVC (% predicted)	76 (71–81)	99 (91–108)
FEV_1/FVC	0.75 (0.72–0.78)	0.77 (0.74–0.80)
Never-smoker (%)	51.2	56.8
Ex-smoker (%)	36.4	35.3
Current smoker (%)	12.4	7.9
Pack-years[#]	23 (13–36)	16 (8–27)
Doctor-diagnosed asthma (%)	16.8	9.9
Doctor-diagnosed COPD (%)	1.7	0.3

Data are presented as mean±SD or median (interquartile range), unless otherwise stated. PRISm: preserved ratio impaired spirometry; BMI: body mass index; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity. [#]: for current and ex-smokers only.

Using the pipeline-derived genetic files, a GWAS of PRISm *versus* controls was performed using BOLT-LMM (https://alkesgroup.broadinstitute.org/BOLT-LMM/BOLT-LMM_manual.html) [15]. The association between PRISm and each SNP was calculated using logistic regression, with SNP coded additively, and adjusting for sex, body mass index (BMI), age and smoking status (smoking status as a dummy variable; 0=never-smoker, 1=ex-smoker, 2=current smoker).

SNPs were filtered to remove those with a MAF ⩽0.01 or that were strand-ambiguous. LD score regression was used to estimate heritability and to assess genomic inflation by calculating λ and LD intercept [16]. To correct for genomic inflation the p-values were corrected for the LD intercept. Stringent LD clumping (r^2=0.001, kb 10 000) was applied to SNPs reaching a significance threshold of $p=5\times10^{-8}$ to define distinct sentinel SNPs. Only SNPs considered novel, based on their reference SNP cluster IDs (rsIDs) not being reported as top signals in the Shrine *et al*. [6] 2019 GWAS of lung function, were investigated in the replicating cohorts. Figure 2 shows a flow chart of the analysis.

Stage 2 and joint analysis

Novel SNPs identified in Stage 1 were tested for association in 13 European ancestry independent cohorts from the SpiroMeta and CHARGE consortia. The supplementary materials summarise full cohort descriptions, spirometry methods, genotyping methods and imputation platforms. Replicating cohorts performed a logistic regression with the lead SNPs from Stage 1 in those with PRISm and control spirometry. Adjustment was made for age, BMI, sex, smoking history (either pack-years or status, as described above) and population substructure by either principal components or using linear mixed models [17]. Results were combined across the Stage 2 studies using a fixed-effect inverse variance model in Stata 17 (StatCorp., College Station, TX, USA). The nearest gene for each SNP was determined using PhenoScanner (www.phenoscanner.medschl.cam.ac.uk).

Definition of top SNPs

We performed a joint analysis of Stage 1 and Stage 2 in a fixed-effect inverse variance model using Stata 17. Heterogeneity was tested for using Cochrane Q. Top SNPs had to meet the following criteria: $p<5\times10^{-8}$ in the joint analysis of Stage 1+2; same direction of effect in Stage 1 and Stage 2; and either p<0.05 in Stage 2 or a lower p-value in the joint Stage 1+2 than in Stage 1. Because no genome-wide significant SNPs have ever been reported in association with PRISm, all top SNPs identified are assumed novel for PRISm.

Genetic correlation

To investigate the shared genetic architecture between PRISm and other traits, we performed a bivariate LD score regression analysis to assess the genome-wide genetic correlation (r_g) between Stage 1 PRISm results and continuous lung function traits, moderate to severe asthma, asthma–COPD overlap, spirometrically diagnosed COPD, respiratory tract infections and eosinophil count [18]. We also examined the genetic correlation between PRISm and related conditions including type 2 diabetes, BMI, hypertension and myocardial infarction.

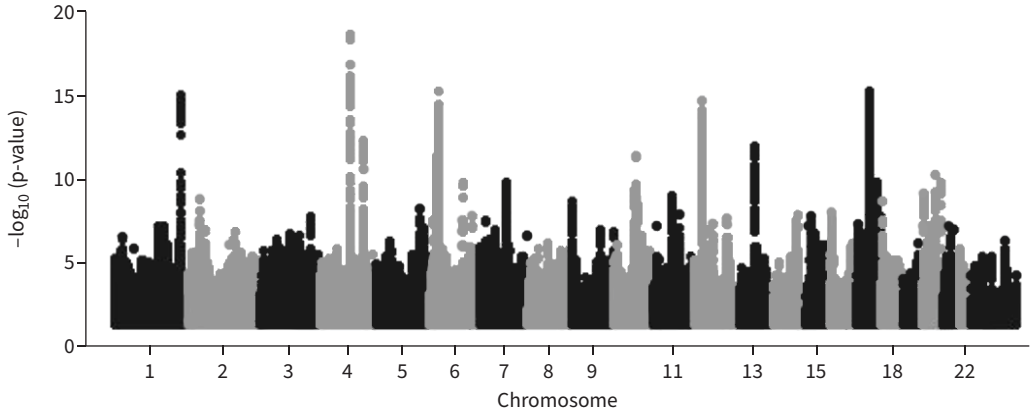

FIGURE 2 Manhattan plot of discovery genome-wide association study after linkeage disequilibrium score regression filtering and adjustment.

PheWAS

We conducted a PheWAS of each top SNP individually using https://gwas.mrcieu.ac.uk/phewas/ to determine if SNP pleiotropy could account for the described associations with PRISm. We highlight lung function measures, lung diseases and comorbidities previously associated with PRISm (*e.g.* overweight or BMI, cardiovascular disease and diabetes) [4].

SNPs novel for lung function: deep PheWAS and gene ontology

Top SNPs that were not reported in the largest GWAS of lung function to date ($r^2 \geqslant 0.5$) were reported as novel for both PRISm and lung function [8]. The SNPs were analysed using deep PheWAS enriched for lung function traits to examine for associations with lung function and other traits [19]. The nearest genes to these novel SNPs were then investigated for gene ontology information using the Functional Mapping and Annotation of Genome-wide association studies (FUMA) tool (https://fuma.ctglab.nl). Hypergeometric tests were implemented using FUMA to test whether the nearest genes to novel SNPs were overrepresented in gene sets from MSigDB, WikiPathways and reported genes from the GWAS catalogue (gene-set enrichment analysis; see https://fuma.ctglab.nl/tutorial#gene2func for full list of gene sets) [20].

Results

Discovery GWAS

FEV_1 and FVC "best measures" were available for 353 315 UKBB participants. Supplementary appendix 1 contains details of spirometric quality control to derive best measure variables. After excluding individuals missing smoking status and/or BMI (n=1441), 38 639 PRISm cases and 257 643 controls were identified.

After further excluding 6490 individuals without derived genetic files in the GWAS pipeline described above, a GWAS of 289 792 individuals was performed (figure 1). A total of 7 339 387 SNPs were tested after exclusions. A Manhattan plot of the results is shown in figure 2. The chip heritability estimate (h2) with SE was 0.0493±0.0024. The λ was 1.25 and the LD intercept 1.02. The 7 339 378 SNP p-values were corrected for the LD intercept, leaving 6037 that met $p<5\times10^{-8}$.

After LD clumping (r^2=0.001, kb 10 000), 33 SNPs from 18 chromosomes remained. We removed SNPs already described in the Shrine *et al.* [6] GWAS, leaving 27 SNPs from 16 chromosomes to investigate in Stage 2.

Stage 2 analysis

The Stage 2 analysis to replicate SNPs discovered in Stage 1 was conducted in 13 cohorts (5165 PRISm cases and 47 729 controls). Stage 2 cohorts used proxies if SNPs were not found in their panel ($r^2 \geqslant 0.8$). SNP rs142330941 was only found in the Framingham cohort and no proxies for it were available in other cohorts, so it was excluded from further analysis, leaving 26 SNPs tested in Stage 2 (figure 3).

In total, 22 SNPs met the criteria for top SNPs. Five of these showed strong evidence of heterogeneity contributing to effect with Cochrane Q $p<0.05$ and $I^2 \geqslant 40\%$. The large difference of sample sizes of Stage 1 and 2 likely contributed to this heterogeneity. Full results at each stage are in table 2.

Genetic correlation

Details of GWAS used for correlation studies with the Stage 1 PRISm can be found in supplementary appendix 2. As expected, we found very strong genetic correlation between PRISm and FEV_1 and FVC, and a gradation of increasing genetic correlation between PRISm and asthma, asthma–COPD overlap and COPD (table 3). Type 2 diabetes showed a moderate genetic correlation with PRISm. Waist-to-hip ratio (after adjustment for BMI) was positively genetically correlated with PRISm, whereas BMI was negatively genetically correlated with PRISm. Cardiac diseases, systolic hypertension and myocardial infarction showed positive genetic correlations with PRISm.

PheWAS

Almost all SNPs had associations with lung function traits (21 of 22 top SNPs). Consistent with the genetic correlation estimates, many SNPs were associated with anthropomorphic traits such as height (12 of 22) and weight or BMI (16 of 22). Associations with diabetes (diagnosis or medication use) or HbA1c (18 of 22) and systolic or diastolic pressure (7 of 22) were common. Full results are in the supplementary tables.

SNPs novel for lung function: deep PheWAS and gene ontology

During the development of this paper, the largest GWAS of lung function was released that describes 1020 SNPs associated with lung function [8]. Four of the PRISm top SNPs were distinct from lung

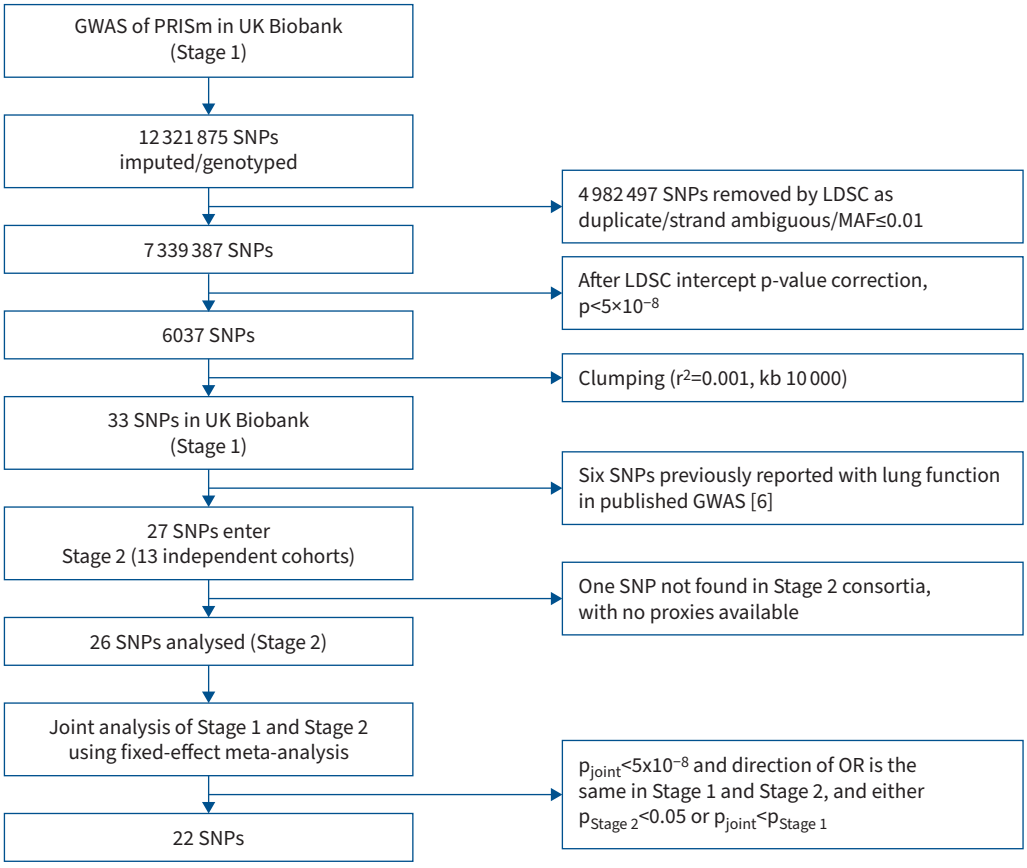

FIGURE 3 Flow chart of single nucleotide polymorphisms (SNPs) analysed. GWAS: genome-wide association study; PRISm: preserved ratio impaired spirometry; LDSC: linkeage disequilibrium score regression; MAF: minor allele frequency; OR: odds ratio.

function signals found in this GWAS (r^2>0.5) and are therefore novel for lung function: rs7652391 (nearest gene *MECOM*), rs9431040 (*HLX*), rs62018863 (*TMEM114*) and rs185937162 (*HLA-B*). We selected these four SNPs for a deep PheWAS enriched for lung function traits. All four SNPs were associated with lung function traits. This is expected because the deep PheWAS used UK Biobank as did our Stage 1 analysis. Rs9431040 was positively associated with systolic blood pressure (supplementary figure E4). As shown in supplementary figure E5, rs185937162 had a strong negative association with inflammatory conditions such as ankylosing spondylitis, uveitis and other spondyloarthropathies. For full results see supplementary tables and figures. Using FUMA, *HLA-B*, *HLX* and, to a lesser extent, *MECOM* were all shown to be expressed in the lung (HLA-B is expressed in nearly all cell types) (supplementary figure E6). *MECOM* encodes a protein that is a transcriptional regulator and oncoprotein and may be involved in haematopoiesis, apoptosis, development and cell differentiation and proliferation. As per the GeneCards database (www.genecards.org), SNPs in the gene have been associated with changes in body height and diastolic blood pressure. *HLX* is predicted to be involved in organ development and is associated with diseases affecting the diaphragm. SNPs in the gene are associated with changes in body height and cholesterol levels. *TMEM114* encodes a protein that has a role in lens and eye development. SNPs in the gene are associated with changes in vital capacity, body height and smoking initiation. *HLA-B* plays a role in the immune system and may influence the susceptibility to infection or the effect of autoimmune processes on the lung. Gene-set enrichment analyses showed that, in addition to respiratory traits, the 26 novel genes were enriched among gene sets for multiple other phenotypes, including white blood cell traits, anthropometric traits and traits relating to diabetes, including β-cell function and glucose levels (supplementary table E12).

Discussion

This is the first GWAS of PRISm to report genetic associations reaching significance thresholds. We show that there is a heritable component for the development of PRISm. We report 22 distinct signals for

https://doi.org/10.1183/13993003.00337-2023

TABLE 2 SNP results Stage 1, Stage 2 and joint Stage 1+Stage 2

rsID	CHR:BP	Nearest gene	Function	EA/NEA	EAF	Stage 1		Stage 2		Stage 1+Stage 2 joint meta-analysis			
						OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	I^2 (%)	Cochrane Q p-value
rs9431040	**1:221152299**	***HLX***	**Intergenic**	**T/C**	**0.28**	**0.99 (0.990–0.994)**	**1.88×10^{-15}**	**0.94 (0.901–0.990)**	**0.020**	**0.99 (0.990–0.994)**	**4.56×10^{-16}**	**0.0**	**0.74**
rs9295	2:36777825	*CRIM1*	UTR3	G/A	0.71	1.01 (1.004–0.008)	2.72×10^{-9}	0.97 (0.923–1.014)	0.175	1.01 (1.004–1.008)	2.61×10^{-9}	3.6	0.41
rs7652391	**3:168913273**	***MECOM***	**Intron**	**G/T**	**0.78**	**0.99 (0.992–0.996)**	**2.29×10^{-8}**	**0.95 (0.902–1.007)**	**0.076**	**0.99 (0.992–0.996)**	**1.10×10^{-8}**	**4.8**	**0.40**
rs6537297	**4:145502029**	***FHDC1***	**Intergenic**	**C/T**	**0.45**	**1.01 (1.005–1.008)**	**7.73×10^{-13}**	**1.05 (1.003–1.093)**	**0.040**	**1.01 (0.996–1.016)**	**2.49×10^{-13}**	**43.7**	**0.03**
rs6923462	**6:7801112**	***BMP6***	**Intron**	**T/C**	**0.84**	**0.99 (0.991–0.996)**	**4.06×10^{-8}**	**0.91 (0.860–0.967)**	**0.002**	**0.99 (0.991–0.995)**	**1.45×10^{-8}**	**40.7**	**0.04**
rs1233604	**6:28734676**	***ZFP57***	**Intergenic**	**G/A**	**0.88**	**0.99 (0.986–0.992)**	**1.18×10^{-15}**	**1.00 (0.929–1.074)**	**0.978**	**0.99 (0.985–0.993)**	**5.96×10^{-16}**	**0.0**	**0.61**
rs185937162	**6:31325268**	***HLA-B***	**Upstream**	**T/G**	**0.96**	**0.99 (0.984–0.992)**	**3.69×10^{-8}**	**0.95 (0.326–2.760)**	**0.339**	**0.99 (0.984–0.992)**	**2.16×10^{-8}**	**13.7**	**0.29**
rs6928024	**6:142551082**	***AL356739.1***	**Intergenic**	**G/A**	**0.75**	**1.01 (1.004–1.008)**	**2.44×10^{-8}**	**1.03 (0.626–1.700)**	**0.226**	**1.01 (1.004–1.008)**	**1.34×10^{-8}**	**39.9**	**0.05**
rs10278266	**7:14943333**	***DGKB***	**Intron**	**A/G**	**0.79**	**0.99 (0.992–0.996)**	**3.49×10^{-8}**	**0.92 (0.875–0.976)**	**0.005**	**0.99 (0.992–0.996)**	**1.36×10^{-8}**	**23.0**	**0.19**
rs9649071	**7:84524701**	***HMGN2P11***	**Intergenic**	**A/G**	**0.75**	**1.01 (1.004–1.008)**	**2.36×10^{-10}**	**1.00 (0.949–1.047)**	**0.915**	**1.01 (1.004–1.008)**	**1.59×10^{-10}**	**23.9**	**0.18**
rs7853305	9:4132402	*GLIS3*	Intron	G/C	0.61	0.99 (0.993–0.996)	1.00×10^{-9}	0.96 (0.924–1.007)	0.106	1.00 (0.993–0.997)	1.46×10^{-9}	6.3	0.38
rs780151	**10:80931481**	***ZMIZ1***	**Intron**	**G/A**	**0.58**	**1.01 (1.003–1.007)**	**5.11×10^{-9}**	**1.07 (1.021–1.113)**	**0.004**	**1.01 (1.003–1.007)**	**1.79×10^{-9}**	**42.5**	**0.03**
rs12808829	**11:62378660**	***EML3***	**Synonymous**	**G/A**	**0.63**	**1.01 (1.004–1.007)**	**1.35×10^{-9}**	**1.05 (1.006–1.101)**	**0.025**	**1.01 (1.004–1.008)**	**5.25×10^{-10}**	**42.5**	**0.03**
rs79487293	12:65905126	*RP11-230G5.2*	Intron	C/T	0.68	0.99 (0.993–0.099)	4.09×10^{-8}	0.98 (0.604–1.611)	0.595	0.99 (0.993–0.997)	4.45×10^{-8}	0.0	0.54
rs11113217	**12:107597518**	***SETP7***	**Intergenic**	**T/C**	**0.38**	**0.99 (0.993–0.997)**	**2.40×10^{-8}**	**0.94 (0.902–0.988)**	**0.012**	**0.99 (0.993–0.997)**	**1.39×10^{-8}**	**24.6**	**0.17**
rs7326916	13:71700945	*LINC00348*	Intron	A/T	0.39	1.01 (1.005–1.008)	9.20×10^{-13}	1.01 (0.648–1.559)	0.817	1.01 (1.004–1.008)	1.01×10^{-12}	0.0	0.52
rs11623779	**14:93096391**	***RIN3***	**Intron**	**T/C**	**0.82**	**1.01 (1.004–1.008)**	**3.58×10^{-8}**	**1.04 (0.587–1.836)**	**0.199**	**1.01 (1.004–1.008)**	**1.95×10^{-8}**	**1.2**	**0.44**
rs11621083	**14:102559538**	***HSP90AA1***	**Intron**	**T/A**	**0.15**	**0.99 (0.991–0.995)**	**1.91×10^{-8}**	**0.97 (0.537–1.740)**	**0.252**	**0.99 (0.991–0.995)**	**1.04×10^{-8}**	**0.0**	**0.56**
rs1717198	**15:41465862**	***EXD1***	**Intergenic**	**T/C**	**0.56**	**1.00 (1.003–1.007)**	**2.72×10^{-8}**	**1.02 (0.665–1.575)**	**0.291**	**1.01 (1.003–1.007)**	**1.54×10^{-8}**	**0.0**	**0.62**
rs2240885	**16:3647098**	***SLX4***	**Intron**	**G/A**	**0.78**	**0.99 (0.992–0.996)**	**1.35×10^{-8}**	**0.94 (0.546–1.611)**	**0.018**	**0.99 (0.992–0.996)**	**5.57×10^{-9}**	**3.1**	**0.42**
rs62018863	**16:8624118**	***TMEM114***	**Upstream**	**G/A**	**0.88**	**0.99 (0.990–0.995)**	**3.10×10^{-8}**	**0.96 (0.479–1.904)**	**0.183**	**0.99 (0.990–0.994)**	**1.65×10^{-8}**	**2.5**	**0.42**
rs139077859	**17:44335579**	***RP11-259G18.3***	**Downstream**	**G/A**	**0.79**	**0.99 (0.989–0.993)**	**1.01×10^{-15}**	**0.93 (0.459–1.896)**	**0.056**	**0.99 (0.989–0.993)**	**3.14×10^{-16}**	**0.0**	**0.50**
rs11651469	**17:69148519**	***CASC17***	**Intron**	**T/G**	**0.44**	**0.99 (0.993–0.996)**	**2.42×10^{-10}**	**0.89 (0.698–0.908)**	**2.6×10^{-23}**	**0.99 (0.992–0.996)**	**1.22×10^{-12}**	**87.1**	**0.00**
rs1000972	**20:6621717**	***RP5-971N18.3***	**Intergenic**	**G/A**	**0.36**	**1.01 (1.004–1.007)**	**1.08×10^{-9}**	**1.04 (0.660–1.625)**	**0.127**	**1.01 (1.004–1.008)**	**4.98×10^{-10}**	**0.0**	**0.56**
rs3091552	**20:45440006**	***AL031055.1***	**Upstream**	**C/G**	**0.27**	**0.99 (0.991–0.995)**	**4.60×10^{-12}**	**0.96 (0.913–1.01)**	**0.121**	**0.99 (0.991–0.995)**	**1.73×10^{-12}**	**20.1**	**0.22**
rs55791529	**20:62363858**	***ZGPAT***	**Intron**	**C/T**	**0.33**	**0.99 (0.992–0.996)**	**2.67×10^{-10}**	**0.95 (0.594–1.521)**	**0.032**	**0.99 (0.992–0.996)**	**1.01×10^{-10}**	**27.0**	**0.15**

Gene was determined using PhenoScanner (https://www.phenoscanner.medschl.cam.ac.uk/). SNPs reaching criteria for top signals ($p_{joint}<5\times10^{-8}$ *and* direction of odds ratio is the same in Stage 1 and Stage 2, *and* either $p_{stage\ 2}<0.05$ *or* $p_{joint}<p_{stage\ 1}$) are highlighted in bold. SNP: single nucleotide polymorphism; CHR:BP: chromosome:base pair position; EA: effect allele; NEA: non-effect allele; EAF: allele 1 frequency.

TABLE 3 Genetic correlation between PRISm and pulmonary and extrapulmonary traits

Trait	$r_g \pm SE$	p-value
FEV_1	−0.96±0.01	<0.001
FVC	−0.93±0.01	<0.001
PEFR	−0.65±0.02	<0.001
FEV_1/FVC	−0.23±0.02	<0.001
COPD	0.62±0.03	<0.001
Asthma–COPD overlap	0.52±0.04	<0.001
Moderate to severe asthma	0.31±0.05	<0.001
Respiratory infection	0.18±0.6	0.003
Blood eosinophils	0.06±0.02	0.012
Type 2 diabetes	0.12±0.03	0.007
BMI	−0.04±0.02	0.031
Waist-to-hip-ratio[#]	0.12±0.02	<0.001
Systolic hypertension	0.08±0.02	<0.001
Diastolic hypertension	0.05±0.02	0.035
Myocardial infarction	0.07±0.03	0.007

PRISm: preserved ratio impaired spirometry; r_g: genome-wide genetic correlation; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity; PEFR: peak expiratory flow rate; BMI: body mass index. [#]: adjusted for BMI.

PRISm that reach genome-wide significance across both stages of the meta-analysis; all represent novel signals for PRISm. Of these, four SNPs were novel for an association with lung function. This demonstrates the usefulness of performing GWAS of different lung function traits and phenotypes to maximise discovery of heritable genetic variants of lung function and disease.

Genetic correlation and PheWAS studies showed there are shared genetic risk factors with other lung function measures and lung conditions such as COPD (r_g=0.62) and asthma–COPD overlap (r_g=0.62), as well as comorbidities of lung disease. In addition to modifiable risk factors such as smoking, COPD can result from a complex interplay of genetic and early life factors that determine lung function trajectories [5, 6]. PRISm is also a heterogeneous state with variable trajectories, of which some progress to COPD over time [2]. Genetic determinants of lung function are associated with COPD [21]. Although not directly tested, given that PRISm and COPD are strongly genetically correlated, this could partially explain the transition between them over time.

PRISm has been consistently associated with systemic comorbidities such as diabetes, heart disease and increased risk of mortality in observational studies [4]. In our analysis, we have shown moderate genetic correlation between PRISm and type 2 diabetes, and the PheWAS showed that 18 of the PRISm SNPs are associated with diabetes, diabetic medication use or hyperglycaemia. PRISm showed positive genetic correlation with type 2 diabetes (r_g=0.12) and waist–hip ratio (r_g=0.12) but a weaker and negative genetic correlation with BMI (r_g= −0.04).

Several observational and cross-sectional studies of impaired lung function (including PRISm) have demonstrated positive associations between PRISm and BMI [7, 22]. Increased BMI has been shown to have stronger associations with restrictive patterns of lung function impairment (such as reduced FVC and FEV_1, but preserved FEV_1/FVC ratio, as seen in PRISm) rather than with the classical obstructive pattern [22]. This could be explained by the mechanical effects of adiposity, where fat accumulates around lungs leading to airway narrowing, and around the abdomen impeding chest wall expansion during full inspiration [22, 23]. However, in our analysis, we observed a slight negative correlation between PRISm and BMI. Notably, in the current GWAS, PRISm was adjusted for BMI, which may have affected the direction and magnitude of the genetic correlation observed.

Other proposed mechanisms for the association between obesity and lung function impairment include systemic inflammation, where levels of pro-inflammatory markers such as interleukin-6, C-reactive protein, fibronectin and other cytokines are increased in people with airway obstruction [22, 24]. The relationship between obesity and impaired lung function may also alter with disease progression. While studies have reported positive correlations between PRISm and adiposity, an inverse relationship is commonly reported between adiposity and COPD, with cachexia being a feature of late-stage COPD. Given that genetic correlations alone cannot imply a direction of causation, comparing childhood- and adult-onset obesity as a

risk factor for PRISm or COPD may help provide insights into the nature and direction of adiposity effects on lung function phenotypes over the life course. Moreover, given the relationship between height as a precursor of maximally attained lung volume, comparing multiple measures of adiposity (including those that do and do not incorporate height) may inform a deeper understanding of relationships between anthropometric traits and lung function.

In our study, various cardiovascular traits showed a degree of genetic correlation with PRISm. Previous observational and genetic correlation studies have shown associations between cardiovascular traits, including high blood pressure and myocardial infarction, and FEV_1 and FVC [24, 25]. There are multiple possible mechanisms for this association, including shared risk factors, such as systemic inflammation induced by cigarette smoking; however, a recent study demonstrated that an association between PRISm and cardiovascular disease (CVD) remained after adjustment for multiple risk factors [26]. Other studies have proposed causal mechanisms for a relationship between PRISm and CVD, including a suggestion that congestive heart failure may lead to lung function impairment as a consequence of cardiomegaly and pulmonary congestion [2]. In terms of shared genetic risk factors, beyond shared specific loci for smoking such as the 15q25.1 locus [27], a study performing partitioned genetic correlation analysis demonstrated genetic correlations between COPD and CVD traits, notably in histone markers (h3k9ac, h3k4me3), which play a central role in arterial pressure and bronchial cell development [27, 28]. Genes such as *HHIP* and *EEFSEC* are known to be associated with lung development signalling pathway and translation factors necessary for protein synthesis associated with COPD and cardiovascular events [27].

We found that there is an overlap between the top SNPs in our PRISm GWAS and continuous lung function traits. This is not surprising given PRISm is a diagnosis based on spirometry, as opposed to a disease defined by a unique pathogenic mechanism. Similarly, a previous GWAS of COPD based on spirometric criteria has discovered loci that have been described as associated with a diverse range of lung diseases, including asthma and idiopathic pulmonary fibrosis [29]. This likely reflects the genetic heterogeneity of lung diseases, as well as how spirometric measures are routinely used for the clinical diagnosis and functional assessment of various lung disorders.

It is possible that better sub-phenotyping of PRISm towards a more homogeneous subpopulation (*e.g.* those with PRISm and BMI $<25\ kg{\cdot}m^{-2}$) could lead to the discovery of new genetic associations. Given that a high proportion of those with PRISm transition to other lung function states over time [4], both normal lung function and COPD, then future genetic studies focusing on persistent PRISm or PRISm that progresses to COPD could be informative, although such focus could result in limited sample size and reduced power.

The majority of published GWAS have adopted a significance threshold of $p<5\times10^{-8}$ [30], although more liberal and stricter thresholds have been proposed [31]. Specifically, stricter thresholds of 3×10^{-8} [32] or 5×10^{-9} [33] have been proposed for GWAS focusing on low frequency (MAF 1–5%) and rare (MAF <1%) genetic variants, respectively, to correct for increased multiple testing. Had we adopted the stricter definition of $p<3\times10^{-8}$, all reported signals would have remained significant. We did not adopt the even stricter threshold of 5×10^{-9} because we excluded rare variants (MAF <1%) from our analysis; however, this may have been at the expense of missing rare variant associations with PRISm.

Population stratification and cryptic relatedness can cause spurious associations in GWAS. We used a linear mixed model for our discovery GWAS, which could account for these issues [34]. To account for genomic inflation we used LD score regression and adjusted for LD intercept. We excluded SNPs that reached significance threshold in Stage 1 analysis if their rsID had been reported as associated with lung function in the largest published lung function GWAS at the time of analysis of Shrine *et al*. [6]. This was in an attempt to focus on SNPs that were novel for PRISm, rather than simple lung function. If they had not been excluded, these SNPs may have been reported in this paper as having an association with PRISm. However, we did not exclude SNPs in LD with previously reported SNPs. Therefore, SNPs in high LD with excluded SNPs may still have been reported. Our GWAS was only performed on those of European heritage; therefore, these results may not be generalisable to other ancestral populations. The SNPs discovered in Stage 1 may not successfully replicate in diverse ancestral populations. The prevalence of PRISm varies by region and ancestral population, from 4.2% in males in Sydney, Australia, to 48.7% in females in Manila, Philippines [35]. Rates of PRISm comorbidities also vary [35]. Future research should aim to recruit from diverse ancestries to explore any heritable component to this variation. Our GWAS was performed using pre-bronchodilator values, although medication was not withheld prior to spirometric testing. Although post-bronchodilator values are not required for PRISm diagnosis, there is evidence that spirometric values can change in those with PRISm post-bronchodilation [36].

Conclusion

This is the first GWAS of PRISm to successfully identify genetic associations reaching genome-wide significance. We defined 22 genetic signals for PRISm, of which four are also novel for lung function, highlighting that GWAS of different lung function phenotypes are complementary. Genetic risk factors for PRISm overlap with those for other lung diseases and extrapulmonary comorbidities.

Acknowledgements: This research has been conducted using the UK Biobank Resource under application number 55521. We acknowledge Marike Boezen's contribution to this paper, who sadly died before its completion. EXCEED is supported by the University of Leicester, the NIHR Leicester Respiratory Biomedical Research Centre, by Wellcome (202849, https://doi.org/10.35802/202849) and by Cohort Access fees from studies funded by the Medical Research Council (MRC), the Biotechnology and Biological Sciences Research Council (BBRSC), National Institute for Health and Care Research (NIHR), the UK Space Agency and GlaxoSmithKline. It was previously supported by MRC grant G0902313. This work is supported by BREATHE – The Health Data Research Hub for Respiratory Health (UKR_PC_19004) in partnership with SAIL Databank. We also thank all participants and staff who have contributed their time to the study. The authors thank all LBC1936 study participants and research team members who have contributed, and continue to contribute, to ongoing studies.

Ethics statement: All LBC1936 participants provided informed consent according to the Declaration of Helsinki and ethical approval was obtained from Multi-Centre Research Ethics Committee for Scotland (MREC/01/0/56) and the Lothian Research Ethics Committee (LREC/2003/2/29).

Data availability: Within four weeks of publication, summary data will be available at https://gwas.mrcieu.ac.uk

Conflict of interest: S.J. London reports support for the present work from National Institutes of Health (NIH), USA. M.H. Cho reports support for the present work from the National Heart, Lung, and Blood Institute (NHLBI) and grants from Bayer outside the submitted work. E. Wan reports grants from US Department of Veterans Affairs Research and Development, outside the submitted work. E. Silverman reports support for the present work from NIH (grant U01 HL089856) and has received institutional grant support from Bayer outside the submitted work. J.D. Crapo reports grants from NIH (U01 HL089897, U01 HL089856); and a leadership role as Chair of Board of Directors of COPD Foundation outside the submitted work. X. Hu reports support for the present manuscript from NIH/NHLBI (R01-HL153248), as well as travel support. L. Lahousse reports consulting fees from AstraZeneca and lecture honoraria from IPSA vzw and Chiesi, outside the submitted work. D.D. Sin reports lecture honoraria from GlaxoSmithKline, AstraZeneca and Boehringer Ingelheim, outside the submitted work. S.E. Harris reports support for the current manuscript from Biotechnology and Biological Sciences Research Council, and the Economic and Social Research Council (BB/W008793/1). M.D. Tobin, A.L. Guyatt and C. John have a funded research collaboration with Orion for collaborative research projects outside the submitted work. G. Davey Smith reports grants from MRC Integrative Epidemiology Unit at the University of Bristol (MC_UU_00011/1) and Scientific Advisory Board Membership for Relation Therapeutics and Insitro, outside the submitted work. M.D. Tobin reports support for the present work from Wellcome Trust, NIHR Leicester Biomedical Research Centre and National Institute of Health and Care Research, and has previously received funding from GlaxoSmithKline for collaborative research projects outside of the submitted work. J.W. Dodd declares he has received honoraria for participating in advisory boards and given lectures at meetings supported by GlaxoSmithKline, Boehringer Ingelheim, Chiesi, AstraZeneca, Fisher & Paykel and Aerogen; he has received sponsorship for attending international scientific meetings from Chiesi; he has also taken part in asthma clinical trials sponsored by Sanofi, AstraZeneca and Chiesi for which his institution received remuneration. All other authors have nothing to disclose.

Support statement: This work was supported by the Medical Research Council (MRC) and the University of Bristol (MC_UU_00011) and MRC CARP Fellowship (MR/T005114/1). J.W. Dodd is supported by the National Institute for Health and Care Research (NIHR) Bristol Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. The Agricultural Lung Health Study (ALHS) was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences (NIEHS) (Z01-ES102385, Z01-ES049030, Z01-ES043012, and for ABW, contract number HHSN273201600003I) and the National Cancer Institute (Z01-CP010119). ALHS was also supported in part by American Recovery and Reinvestment Act (ARRA) funds through NIEHS contract number N01-ES-55546. The ALHS thank the numerous study staff at Social and Scientific Systems, Inc. who played a role in the data collection, and acknowledge Jianping Jin (Westat) for expert computational assistance on the ALHS and Nathan Gaddis (RTI) for assistance performing genotype QC and HRC imputation for the ALHS. The research was partially supported by the NIHR Leicester Biomedical Research Centre and through an NIHR Senior Investigator Award to M.D. Tobin; views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in the design of the study. This research was funded in whole, or in part, by the Wellcome Trust: Wellcome Trust Investigator Award (WT202849/Z/16/Z, M.D. Tobin) and Wellcome

Trust Discovery Award (WT225221/Z/22/Z, M.D. Tobin and A.L. Guyatt). MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Analyses for this project were supported by NIH/NHLBI R01-HL153248. The MESA Lung Study is supported by R01-HL077612 and R01-HL093081. Support for MESA is provided by contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001420, UL1-TR-001881 and DK063491. This publication was developed under Science to Achieve Results (STAR) research assistance agreements, No. RD831697 (MESA Air) and RD-83830001 (MESA Air Next Stage), awarded by the US Environmental Protection Agency (EPA). It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication. Funding for SHARe genotyping was provided by NHLBI contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, CA, USA) and the Broad Institute of Harvard and MIT (Boston, MA, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0. The authors acknowledge Research Computing at The University of Virginia for providing computational resources and technical support that have contributed to the results reported within this publication (https://rc.virginia.edu). J.V.T. Lominchar, T. Hansen and N. Grarup were funded by the Novo Nordisk Foundation (grant number NNF18CC0034900). Cardiovascular Health Study: This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086 and 75N92021D00006; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393 and U01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, Clinical and Translational Science Institute grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. LBC1936 is supported by the Biotechnology and Biological Sciences Research Council (BBSRC), and the Economic and Social Research Council (BB/W008793/1) (which supports S.E. Harris), Age UK (Disconnected Mind project) and the University of Edinburgh. S.R. Cox is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (221890/Z/20/Z). Genotyping of the cohorts was funded by the BBSRC (BB/F019394/1). The Rotterdam Study is supported by the Erasmus MC and Erasmus University Rotterdam; the Netherlands Organisation for Scientific Research (NWO) (grant 918-46-615); the Netherlands Organisation for Health Research and Development (ZonMW); the Research Institute for Diseases in the Elderly (RIDE); the Netherlands Genomics Initiative; the Ministry of Education, Culture and Science; the Ministry of Health, Welfare and Sports; the European Commission (DG XII; grant 601055); and the Municipality of Rotterdam. S.R.A. Wijnant was funded by the “Funds for Scientific Research Flanders (Fonds voor Wetenschappelijk Onderzoek Vlaanderen)”, grant 3G037618. The COPDGene project was supported by award number U01 HL089897 and award number U01 HL089856 from the NHLBI. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NHLBI or the NIH. COPDGene is also supported by the COPD Foundation through contributions made to an Industry Advisory Board that has included AstraZeneca, Bayer Pharmaceuticals, Boehringer Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer and Sunovion. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Stefano G, Duane LS, Claire V, *et al.* Morbidity and mortality associated with the restrictive spirometric pattern: a longitudinal study. *Thorax* 2010; 65: 499–504.
2 Wijnant SRA, De Roos E, Kavousi M, *et al.* Trajectory and mortality of preserved ratio impaired spirometry: the Rotterdam Study. *Eur Respir J* 2020; 55: 1901217.
3 Wan ES, Fortis S, Regan EA, *et al.* Longitudinal phenotypes and mortality in preserved ratio impaired spirometry in the COPDGene study. *Am J Respir Crit Care Med* 2018; 198: 1397–1405.
4 Higbee DH, Granell R, Davey Smith G, *et al.* Prevalence, risk factors, and clinical implications of preserved ratio impaired spirometry: a UK Biobank cohort analysis. *Lancet Respir Med* 2021; 10: 149–157.
5 Agusti A, Noell G, Brugada J, *et al.* Lung function in early adulthood and health in later life: a transgenerational cohort analysis. *Lancet Respir Med* 2017; 5: 935–945.
6 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: 481–493.
7 Wan ES, Castaldi PJ, Cho MH, *et al.* Epidemiology, genetics, and subtyping of preserved ratio impaired spirometry (PRISm) in COPDGene. *Respir Res* 2014; 15: 89.

8 Shrine N, Izquierdo AG, Chen J, *et al.* Multi-ancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. *Nat Genet* 2023; 55: 410–422.
9 Bycroft C, Freeman C, Petkova D, *et al.* Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv* 2017; preprint [https://doi.org/10.1101/166298].
10 IEU. MRC IEU UK Biobank GWAS Pipeline Version 2. 2019. Available from: https://data.bris.ac.uk/data/dataset/pnoat8cxo0u52p6ynfaekeigi
11 Mitchell RE, Hemani G, Dudding T, *et al.* UK BioBank Genetic Data: MRC-IEU Quality Control, version 2. Bristol, University of Bristol, 2019.
12 McCarthy S, Das S, Kretzschmar W, *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016; 48: 1279–1283.
13 O'Connell J, Sharp K, Shrine N, *et al.* Haplotype estimation for biobank-scale data sets. *Nat Genet* 2016; 48: 817–820.
14 Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* 2011; 1: 457–470.
15 Loh PR, Kichaev G, Gazal S, *et al.* Mixed-model association for biobank-scale datasets. *Nat Genet* 2018; 50: 906–908.
16 Bulik-Sullivan BK, Loh PR, Finucane HK, *et al.* LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; 47: 291–295.
17 Price AL, Patterson NJ, Plenge RM, *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; 38: 904–909.
18 Bulik-Sullivan B, Finucane HK, Anttila V, *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015; 47: 1236–1241.
19 Packer RJ, Williams AT, Hennah W, *et al.* DeepPheWAS: an R package for phenotype generation and association analysis for phenome-wide association studies. *Bioinformatics* 2023; 39: btad073.
20 Watanabe K, Taskesen E, van Bochoven A, *et al.* Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017; 8: 1826.
21 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.
22 Wielscher M, Amaral AFS, van der Plaat D, *et al.* Genetic correlation and causal relationships between cardio-metabolic traits and lung function impairment. *Genome Med* 2021; 13: 104.
23 Zhu Z, Li J, Si J, *et al.* A large-scale genome-wide association analysis of lung function in the Chinese population identifies novel loci and highlights shared genetic aetiology with obesity. *Eur Respir J* 2021; 58: 2100199.
24 Li D, Ruan Z, Xie S, *et al.* The relationship between preserved ratio impaired spirometry and mortality in the myocardial infarction survivors: a population-based cohort study. *BMC Cardiovasc Disord* 2023: 23: 331.
25 Schnabel E, Karrasch S, Schulz H, *et al.* High blood pressure, antihypertensive medication and lung function in a general adult population. *Respir Res* 2011; 12: 50.
26 Krishnan S, Tan WC, Farias R, *et al.* Impaired spirometry and COPD increase the risk of cardiovascular disease: a Canadian cohort study. *Chest* 2023; 164: 637–649.
27 Zhu Z, Wang X, Li X, *et al.* Genetic overlap of chronic obstructive pulmonary disease and cardiovascular disease-related traits: a large-scale genome-wide cross-trait analysis. *Respir Res* 2019; 20: 64.
28 Zhu Z, Hasegawa K, Camargo CA, *et al.* Investigating asthma heterogeneity through shared and distinct genetics: insights from genome-wide cross-trait analysis. *J Allergy Clin Immunol* 2020; 147: 796–807.
29 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.
30 McCarthy MI, Abecasis GR, Cardon LR, *et al.* Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008; 9: 356–369.
31 Chen Z, Boehnke M, Wen X, *et al.* Revisiting the genome-wide significance threshold for common variant GWAS. *G3* 2021; 11: jkaa056.
32 Fadista J, Manning AK, Florez JC, *et al.* The (in)famous GWAS p-value threshold revisited and updated for low-frequency variants. *Eur J Hum Genet* 2016; 24: 1202–1205.
33 Pulit SL, de With SA, de Bakker PI. Resetting the bar: statistical significance in whole-genome sequencing-based association studies of global populations. *Genet Epidemiol* 2017; 41: 145–151.
34 Loh P-R, Tucker G, Bulik-Sullivan BK, *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* 2015; 47: 284–290.
35 Mannino DM, McBurnie MA, Tan W, *et al.* Restricted spirometry in the Burden of Lung Disease Study. *Int J Tuberc Lung Dis* 2012; 16: 1405–1411.
36 Backman H, Eriksson B, Hedman L, *et al.* Restrictive spirometric pattern in the general adult population: methods of defining the condition and consequences on prevalence. *Respir Med* 2016; 120: 116–123.