Human Molecular Genetics, 2021, Vol. 30, No. 7 619–628

doi: 10.1093/hmg/ddab068
Advance Access Publication Date: 10 March 2021
Association Studies Article

ASSOCIATION STUDIES ARTICLE

Fate or coincidence: do COPD and major depression share genetic risk factors?

Victoria L. Martucci[1,2], Bradley Richmond[3,4], Lea K. Davis[1,2,5], Timothy S. Blackwell[3,4,6], Nancy J. Cox[1,2], David Samuels[1,7], Digna Velez Edwards[1,8,9] and Melinda C. Aldrich[1,2,9,10,11,*]

[1]Vanderbilt Genetics Institute, Vanderbilt University School of Medicine, Nashville, TN 37232, USA, [2]Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, USA, [3]Department of Veterans Affairs Medical Center, Nashville, TN 37212, USA, [4]Division of Allergy, Pulmonary, and Critical Care Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, USA, [5]Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN 37232, USA, [6]Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN 37232, USA, [7]Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA, [8]Division of Quantitative Sciences, Department of Obstetrics and Gynecology, Vanderbilt University Medical Center, Nashville, TN 37232, USA, [9]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN 37232, USA, [10]Department of Thoracic Surgery, Vanderbilt University Medical Center, Nashville, TN 37232, USA and [11]Division of Epidemiology, Vanderbilt University Medical Center, Nashville, TN 37232, USA

*To whom correspondence should be addressed at: Vanderbilt University Medical Center, 519A Light Hall, 2215 Garland Avenue, Nashville, TN 37232, USA. Tel: 615 8752627; Fax: 615 9362510; Email: melinda.aldrich@vumc.org

Abstract

Major depressive disorder (MDD) is a common comorbidity in chronic obstructive pulmonary disease (COPD), affecting up to 57% of patients with COPD. Although the comorbidity of COPD and MDD is well established, the causal relationship between these two diseases is unclear. A large-scale electronic health record clinical biobank and genome-wide association study summary statistics for MDD and lung function traits were used to investigate potential shared underlying genetic susceptibility between COPD and MDD. Linkage disequilibrium score regression was used to estimate genetic correlation between phenotypes. Polygenic risk scores (PRS) for MDD and lung function traits were developed and used to perform a phenome-wide association study (PheWAS). Multi-trait-based conditional and joint analysis identified single-nucleotide polymorphisms (SNPs) influencing both lung function and MDD. We found genetic correlations between MDD and all lung function traits were small and not statistically significant. A PRS–MDD was significantly associated with an increased risk of COPD in a PheWAS [odds ratio (OR) = 1.12, 95% confidence interval (CI): 1.09–1.16] when adjusting for age, sex and genetic ancestry, but this relationship became attenuated when controlling for smoking history (OR = 1.08, 95% CI: 1.04–1.13). No significant associations were found between the lung function PRS and MDD. Multi-trait-based conditional and joint analysis identified three SNPs that may contribute to both traits, two of which were previously associated with mood disorders and COPD. Our findings suggest that the observed relationship between COPD and MDD may not be driven by a strong shared genetic architecture.

Received: August 28, 2020. **Revised:** February 24, 2021. **Accepted:** February 27, 2021

Introduction

Chronic obstructive pulmonary disease (COPD) is a leading cause of morbidity and mortality globally, affecting 328 million people and causing 3 million deaths per year (1). Comorbidities are common among COPD patients (2,3). Individuals with comorbid conditions report decreased quality of life (4–7), and the presence of multiple comorbidities can increase mortality rates by as much as 400% (8). Therefore, understanding the relationship between COPD and its comorbidities is a research priority (9).

Psychiatric comorbidities are commonly reported in COPD patients. Individuals with COPD have an increased prevalence of major depression, with estimates ranging from 8 to 80% (10–15). The prevalence of depression is higher in individuals with more severe disease (10,11,14). Among individuals with COPD, depression is associated with greater exacerbation, higher rates of hospital re-admission, decreased medication adherence, poorer quality of life and increased mortality (10–12,16–21).

The biologic mechanism underlying the relationship between COPD and depression is unknown. Both disorders are highly heritable, with an estimated genetic heritability of 25–37% for COPD (22) and 28–51% for major depressive disorder (MDD) (23–25). Heritability of lung function traits such as forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC), which are the basis for COPD diagnosis, are also high, with estimated heritability ranging from 18 to 50% (26–28). Systemic inflammation, hypoxemia and oxidative stress, and shared environmental risk factors, such as smoking, have been proposed as possible mechanisms linking these two conditions (12,29–31). Smoking is a major risk factor for COPD, and it may also be an independent risk factor for depression, though the direction of this relationship is still debated (31–33). Shared genetic risk factors have been investigated in a small number of studies (34–37). A candidate gene study for depression identified a small number of single-nucleotide polymorphisms (SNPs) associated with increased COPD risk (34). A large-scale phenome-wide association study (PheWAS) conducted in the UK Biobank detected associations between lung function genomic loci and depressive symptoms (35). These studies suggest that the relationship between COPD and MDD may be due to pleiotropy, where a single SNP affects two or more distinct traits (38). However, a genome-wide association study (GWAS) of depressive symptoms in smokers with COPD did not identify any significant loci (36). A polygenic risk score (PRS) built from a genome-wide gene-by-environment interaction study of depressive symptoms identified a significant association with COPD, but the underlying model assumed an interaction between SNPs and stressful life events and therefore did not examine purely genetic effects (37). Further complicating the relationship between COPD and MDD is the presence of sex differences in both disorders. MDD is more prevalent in women, and women typically experience more severe depressive symptoms than men (39). Genetic studies of MDD have identified evidence of sex-specific risk variants and transcriptional signatures (40,41). Women develop COPD at lower smoke exposure than men and may experience more severe disease and rapid respiratory decline compared with men with similar smoking exposure (42–44). We investigated the genetic relationship between COPD and MDD, using existing GWAS summary statistics to test for genetic correlation and pleiotropy between the traits. We leveraged electronic health records (EHR) linked to genotyping data to explore shared genetic associations between COPD and MDD using a PheWAS, an approach often used to examine relationships between comorbid conditions (45–47). We also performed sex-stratified analyses to investigate possible sex differences in the relationship between MDD and COPD. An overall schematic of our study design and methods is provided in Supplementary Material, Figure S1.

Results

Study population

Our BioVU study population consisted of 72 447 European ancestry individuals with 9 386 383 SNPs. Approximately 5% of the BioVU population had a COPD phecode. COPD individuals were older (median age 68 years) and more male (53.5%) than the overall study population (median age 56 years and 44.0% male). COPD patients had a higher prevalence of ever smoking (87.6%) than the overall BioVU population (49.8%). The prevalence of major depression (one or more depression phecodes) was higher in COPD patients (8.8%) than among patients without a diagnosis of COPD (3.5%) (Table 1).

Genetic correlation between MDD and lung function

We found low genetic correlations (*Rg*) between MDD and lung function traits using linkage disequilibrium score regression (LDSC). None of the genetic correlations between MDD and lung function were statistically significant. The strongest correlation between MDD and lung function was with peak expiratory flow (PEF) ($Rg = -0.035$, $P = 0.07$). In contrast, we observed strong and statistically significant correlation between lung function traits (Table 2). Local genetic correlation showed statistically significant peaks in *Rg* on chromosome 6 for both FEV_1/FVC (Bonferroni-corrected *P*-value $= 8.62 \times 10^{-3}$) and FEV_1 and MDD (Bonferroni-corrected *P*-value $= 4.38 \times 10^{-6}$). However, the maximum correlation values were still small (maximum *Rg* for FEV_1/FVC and MDD: 3.86×10^{-4}, maximum *Rg* for FEV_1 and MDD: 3.36×10^{-4}).

PheWAS analyses with lung function and MDD–PRS

We built PRS for lung function (818 738 SNPs) and MDD (803 205 SNPs) from publicly available GWAS summary statistics for lung function measures and MDD. To confirm expected associations with lung function, we used linear regression to test for the association between the lung function PRS and their corresponding pre-bronchodilator lung function traits in a subset of BioVU patients with available lung function data. The PRS were robustly associated with the corresponding lung function traits (data not shown). We performed a PheWAS using logistic regression models to examine associations between PRSs and 1857 phecodes in the entire study population. Cases and controls were defined independently for each phecode, and phecodes with <20 cases were excluded ($N = 438$ phecodes). The lung function PRS were consistently associated with decreased COPD in the PheWAS (Table 3). Similar associations were observed in sex-stratified analyses, though the significance of the association varied between lung function phenotypes (Supplementary Material, Figs S2–S5). The MDD–PRS was significantly associated with increased risk of mood disorders [odds ratio (OR) = 1.28, 95% confidence interval (CI): 1.25–1.32; $P = 6.42 \times 10^{-76}$] and MDD (OR = 1.27, 95% CI: 1.22–1.32; $P = 1.41 \times 10^{-31}$) when adjusting for age, sex and the first three principal components (PCs) (Table 3). In sex-stratified analyses, the MDD–PRS was also significantly

Table 1. Demographics of European ancestry BioVU population (2007–2019)

Characteristic	COPD phecode ($N = 3466$)	No COPD phecode ($N = 68\,981$)	Total ($N = 72\,447$)
Median age (IQR)	68 (60–76)	55 (35–68)	56 (36–68)
Gender (*N*, %)			
Female	1615 (46.6)	38 969 (56.5)	40 584 (56.0)
Male	1851 (53.4)	30 010 (43.5)	31 861 (44.0)
Missing	0	2	2
Smoking status (*N*, %)			
Ever	2435 (83.9)	20 861 (41.2)	23 296 (43.5)
Never	455 (16.1)	29 741 (58.8)	30 207 (56.5)
Missing	565	18 379	18 944
Major depressive disorder (*N*, %)	305 (8.8)	2385 (3.5)	2690 (3.7)

COPD, chronic obstructive pulmonary disease

Table 2. Genetic correlation between major depressive disorder and lung function traits

Phenotype 1	Phenotype 2	*Rg*	*P*-value
MDD	FEV_1/FVC	−0.0011	0.95
MDD	FEV_1	−0.0325	0.07
MDD	FVC	−0.0307	0.10
MDD	PEF	−0.0351	0.07
FEV_1/FVC	FEV_1	0.4046	2.66×10^{-89}
FEV_1/FVC	FVC	−0.0841	3.20×10^{-5}
FEV_1/FVC	PEF	0.6273	0
FEV_1	FVC	0.877	0
FEV_1	PEF	0.7058	0
FVC	PEF	0.4351	1.28×10^{-136}

FEV_1, forced expiratory volume in one second; FVC, forced vital capacity; MDD, major depressive disorder; PEF, peak expiratory flow.

associated with mood disorders and MDD (Supplementary Material, Fig. S6, Supplementary Material, Table S1).

In addition to the expected phenotype associations, we observed a significant association between the MDD–PRS and COPD when adjusting for age, sex and the first three PCs (OR = 1.13; 95% CI: 1.09–1.17; *P*-value $=3.72 \times 10^{-12}$) (Table 3, Fig. 1A). Adjusting for smoking attenuated the association and was no longer statistically significant (OR = 1.09; 95% CI: 1.04–1.13; $P = 8.07 \times 10^{-5}$) (Table 3, Fig. 1B). Similar patterns were observed for both men and women in the sex-stratified analyses of PRS–MDD (Supplementary Material, Table S1, Supplementary Material, Fig. S6). None of the lung function PRS were associated with MDD in the smoking-adjusted or smoking-unadjusted analyses (Table 3, Fig. 2). Similarly, no significant associations between any of the lung function PRS and MDD were observed in the sex-stratified analyses (Supplementary Material, Table S1, Supplementary Material, Figs S2–S5).

Multi-trait conditional analysis to detect potential pleiotropy

We used multi-trait-based conditional and joint analysis (mtCOJO) to adjust MDD for the genetic effects of FEV_1/FVC. The majority of SNPs showed little to no change in the effect estimate. The median percent change in the beta before and after conditioning was 0%, with an inter-quartile range of −6–5%. However, heterogeneity in dependent instrument outlier approach (HEIDI-outlier) identified three SNPs (rs12040241, rs7617480, rs12967855) with evidence of pleiotropy between MDD and FEV_1/FVC (Supplementary Material, Table S2).

Discussion

We evaluated the potential for shared genetic architecture between lung function and MDD. We did not observe a significant global genetic correlation between lung function traits and MDD, consistent with prior work (48). In contrast, genetic correlations between lung function traits ranged from −0.08 to 0.87, similar to previous studies (35). Local genetic correlation did identify a small but statistically significant increase in genetic correlation on chromosome 6 in the human leukocyte antigen region. This finding is consistent with the known role of inflammation and the immune system in both COPD (49,50) and MDD (51,52). We found that the PRS–MDD was significantly associated with COPD in our PheWAS, but this association was no longer statistically significant when controlling for smoking. Conversely, none of the lung function PRS showed a significant association with MDD in PheWAS analyses, suggesting little shared genetic architecture between lung function and MDD. However, using multi-trait conditional analysis, we identified three potentially pleiotropic SNPs. Interestingly, two of these SNPs were associated with both mood and smoking traits in a prior GWAS (53–58). An intronic variant in *KLHDC8B*, rs7617480, was previously identified as genome-wide significant in GWAS of smoking cessation (53) and subjective well-being (54). The second SNP, rs12967855, an intronic variant in *CELF4*, was previously found to have genome-wide significant associations with lifetime smoking index (55) and unipolar depression (56–58).

Although we identified three potentially pleiotropic variants, our findings do not provide strong evidence for a shared genetic architecture between MDD and COPD. Smoking behaviors may contribute to the relationship between MDD and COPD (32,33,59). Cigarette smoking and nicotine dependence have been identified as potential confounding factors of the relationship between COPD and mood disorders (59), and smoking may modify associations between COPD and depression (31). Among individuals with COPD, current smokers report higher rates of depression symptoms and have increased mortality risks compared with former smokers and individuals without depression (60,61). Previous studies have also shown that smokers with mental illness have higher mortality rates, particularly from respiratory conditions (60,62–64). Further study is needed to understand

Table 3. Association of lung function and MDD–PRS with COPD and MDD in European BioVU participants (2007–2019)

PRS	COPD				MDD			
	OR[a]	95% CI[a]	OR[b]	95% CI[b]	OR[a]	95% CI[a]	OR[b]	95% CI[b]
FEV_1	0.87	0.84–0.90	0.87	0.84–0.90	1.00	0.96–1.04	0.99	0.95–1.03
FVC	0.94	0.91–0.98	0.95	0.91–0.99	1.00	0.96–1.04	0.99	0.95–1.03
FEV_1/FVC	0.83	0.81–0.86	0.83	0.80–0.87	1.00	0.96–1.04	1.01	0.97–1.05
PEF	0.89	0.86–0.92	0.88	0.85–0.92	1.03	0.99–1.07	1.03	0.99–1.07
MDD	1.13	1.09–1.17	1.07	1.03–1.12	1.27	1.22–1.32	1.24	1.19–1.30

COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in one second; FVC, forced vital capacity; MDD, major depressive disorder; PEF, peak expiratory flow; OR, odds ratio; CI, confidence interval.
[a]Model adjusted for age, sex and first three principal components (N = 72 445).
[b]Model adjusted for age, sex, first three principal components and ever smoking (N = 53 503).

the underlying mechanisms linking smoking, COPD and MDD (59–61).

This study has several strengths and considerations. We used available summary statistics from large, well-powered GWAS to conduct our analyses (35,65,66). We also used the rich BioVU resource with extensive clinical data allowing us to examine multiple phenotypes. Our study is limited by the inclusion of only European ancestry participants. PRS performance decreases in cross-ancestry analysis (67,68), and the limited number of lung function GWAS that have been conducted in African Americans have had small sample sizes with few genome-wide significant findings (69,70). Further research is needed to understand the genetic relationship between COPD and MDD in non-European descent populations. Another limitation of our study is the lack of a replication population to validate our findings. However, our findings are consistent with prior research (35,48). Finally, our study relied on EHR data, which can present challenges due to data missingness and misclassification (71–75). We chose to use phecodes to define phenotypes in our study, as previous research has demonstrated that phecodes better capture clinical disease than International Classification of Disease (ICD) codes alone (76). For the majority of phenotypes, we expect the effects of misclassification to be minimal or biased toward the null (77,78). We also encountered challenges due to missingness, particularly for smoking data (Table 1), which is prone to high rates of missingness and inaccuracies in EHR (79–82). Individuals who were missing smoking data were younger and had a lower prevalence of COPD than those with available smoking information (Supplementary Material, Table S3), thus relying on complete case analysis may limit the generalizability of our findings.

In conclusion, we found that the elevated prevalence of MDD in COPD cannot be solely explained by shared genetic risk factors. Our findings suggest a role for shared environmental or behavioral risk factors, such as smoking. We identified three potentially pleiotropic SNPs that can be prioritized in future studies of MDD and COPD. These findings require further investigation into the biological underpinnings between MDD and COPD to elucidate the causal mechanism underlying their relationship.

Materials and Methods

Study population

Our study population included participants in the Vanderbilt University Medical Center BioVU clinical repository (2007–2019). BioVU is a DNA biobank linked to de-identified EHR clinical data, dating back to the 1980s (83). We limited our study population to BioVU individuals of European ancestry previously genotyped on the Illumina Infinium Multi-Ethnic Genotyping Array. Demographic data (sex, age at last record), smoking, ICD-9 and ICD-10 codes, and pulmonary function data (2011–2019) were extracted from structured fields in the EHR using natural language processing.

We selected individuals of European ancestry using PC analysis implemented in EIGENSTRAT (84,85). We performed standard quality control and imputed genotypes to the Haplotype Reference Consortium with the Michigan Imputation Server (86). Genotypes were hard-called using default settings ($P > 0.1$) in PLINK 1.9 (87,88).

GWAS summary statistics

To investigate potential pleiotropy between lung function and MDD, we used publicly available summary statistics from previously performed GWAS in individuals of European ancestry. Summary statistics were obtained from a large-scale GWAS of lung function (FEV_1, FVC, FEV_1/FVC and PEF) (35) and from a meta-analysis of two genome-wide studies of MDD (65,66).

Genetic correlation

We calculated the overall *Rg* between traits using LDSC software and a reference linkage disequilibrium (LD) score panel derived from European 1000 Genomes populations (89,90). To calculate local genetic correlation, we used Heritability Estimation from Summary Statistics (ρ-HESS) with a European LD reference panel provided by the software authors (91).

Polygenic risk scores

To build PRS, we used polygenic risk score-continuous shrinkage to estimate posterior effect sizes of SNPs with continuous shrinkage priors in each GWAS (92). We then applied the score function in PLINK 1.9 (87,88) to calculate a PRS for each individual in BioVU. PRS were normalized by subtracting the mean and dividing by the standard deviation.

PheWAS

We explored the relationship between each PRS and EHR phenotypes in a PheWAS (93). We performed logistic regression analysis to examine associations between PRS and 1857 phecodes. Phecodes are defined by aggregating similar ICD-9 and ICD-10 billing codes (76,94) and have been used extensively in prior studies (95–108). We mapped extracted ICD-9 and ICD-10

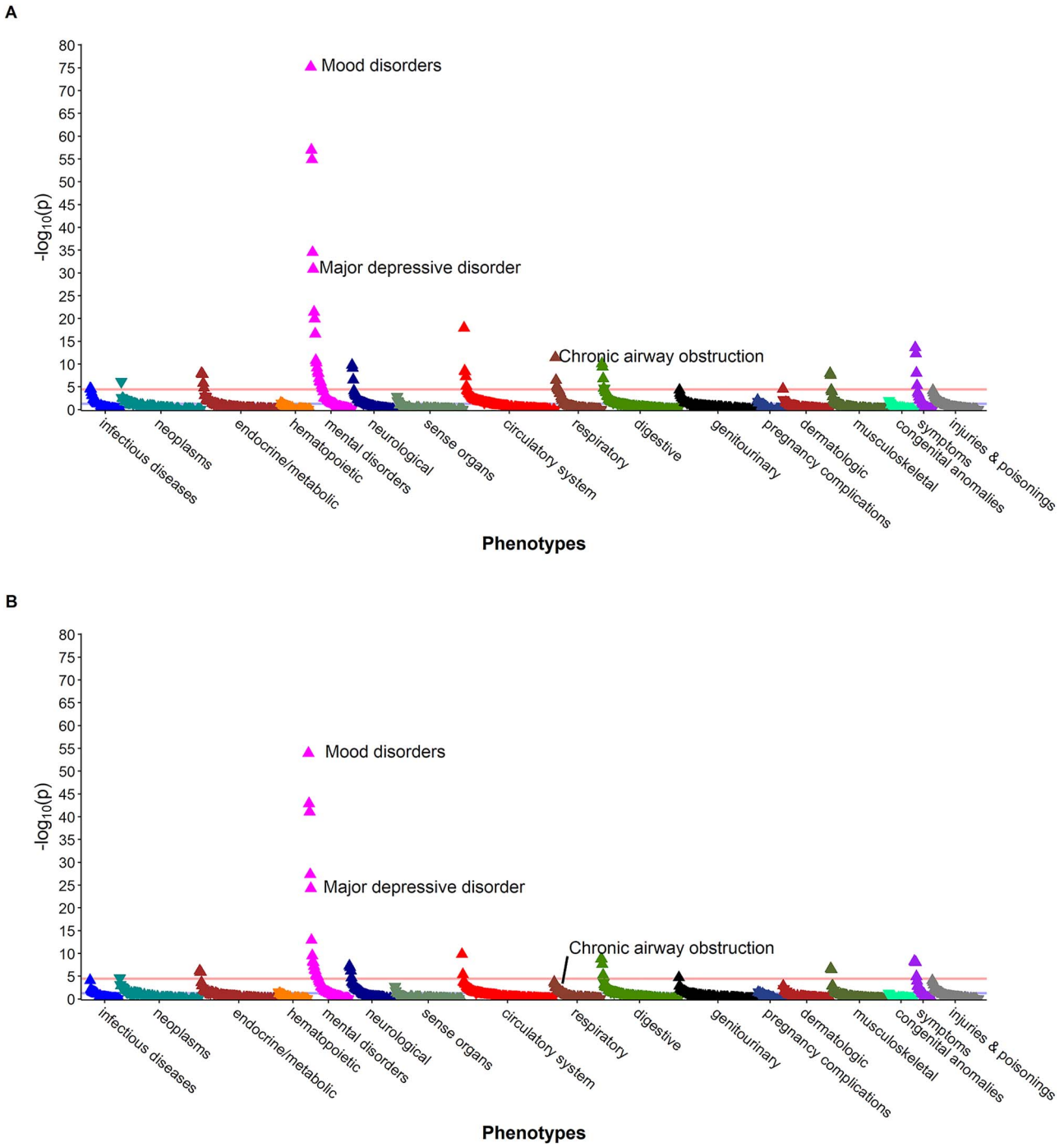

Figure 1. Phenome-wide association study among BioVU participants (2007–2019) of major depression polygenic score, adjusted for (**A**) age, sex and first three principal components, and (**B**) age, sex, first three principal components and ever smoking.

billing codes from BioVU to phecodes using the PheWAS R package (109). Phecodes with fewer than 20 cases were excluded from analyses. Models were adjusted for age at last visit, sex, smoking (ever/never) and three PCs estimated using EIGENSTRAT (84,85) to adjust for potential confounding by genetic ancestry. We also performed sex-stratified PheWAS using the same parameters and covariates as in the main analysis, with the exception of sex as a covariate. A type 1 error rate of alpha = 0.05/1857 phecodes = 2.69×10^{-5} was set for inference of statistical significance.

Multi-trait conditional analysis

We performed mtCOJO to investigate cross-phenotype effects (110). We evaluated the change in effect size for SNPs in the FEV_1/FVC GWAS before and after conditioning on MDD. We also implemented HEIDI-outlier, incorporated into mtCOJO methods, to detect potentially pleiotropic SNPs (110). We used the National Human Genome Research Institute-European Bioinformatics Institute (NHGRI-EBI) GWAS Catalog (111) and the National Heart, Lung, and Blood Institute (NHLBI) Genome-Wide

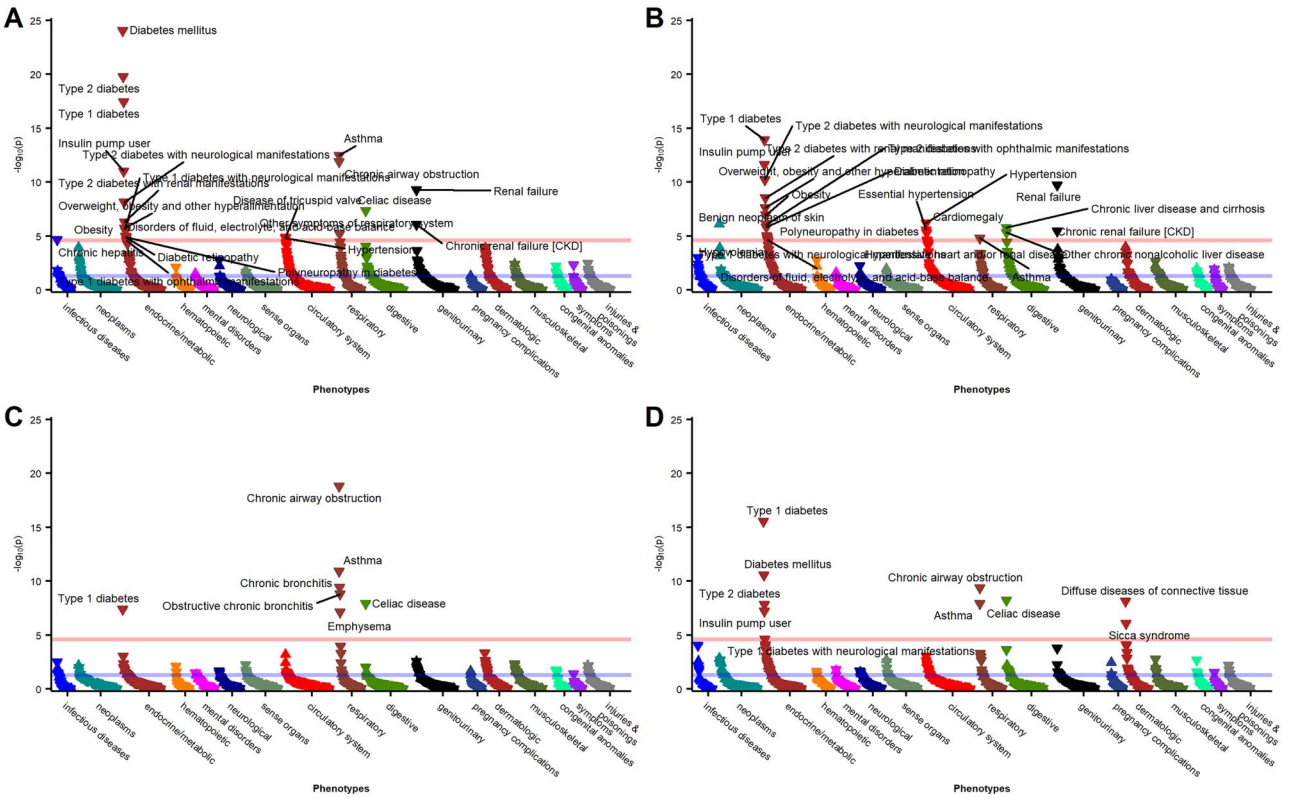

Figure 2. Phenome-wide association study among BioVU participants (2007–2019) of (**A**) FEV_1, (**B**) FVC, (**C**) FEV_1/FVC and (**D**) PEF polygenic scores, adjusted for age, sex, first three principal components and ever smoking.

Repository of Associations Between SNPs and Phenotypes (112) to look up prior associations for identified SNPs.

Supplementary Material

Supplementary material is available at *HMG* online.

Acknowledgement

The authors would like to thank Peter Straub for performing quality control and imputation of the genotyping data.

Conflict of Interest statement. The authors declare no conflict of interest.

Funding

National Institute of General Medical Sciences (T32GM007347, T32GM080178); National Heart, Lung, and Blood Institute (F30HL140756); Department of Veterans Affairs (Grant 5 IK2 BX 003841-02); National Center for Advancing Translational Sciences Clinical and Translational Science Awards (UL1TR002243); National Institute of Mental Health (1R56MH120736-01).

References

1. World Health Organization. (2018) *Global Health Estimates 2016: Disease Burden by Cause, Age, Sex, by Country and by Region, 2000-2016*. World Health Organization, Geneva.
2. Putcha, N., Drummond, M.B., Wise, R.A. and Hansel, N.N. (2015) Comorbidities and chronic obstructive pulmonary disease: prevalence, influence on outcomes, and management. *Semin. Respir. Crit. Care Med.*, **36**, 575–591.
3. Baty, F., Putora, P.M., Isenring, B., Blum, T. and Brutsche, M. (2013) Comorbidities and burden of COPD: a population based case-control study. *PLoS One*, **8**, e63285.
4. Putcha, N., Puhan, M.A., Hansel, N.N., Drummond, M.B. and Boyd, C.M. (2013) Impact of co-morbidities on self-rated health in self-reported COPD: an analysis of NHANES 2001–2008. *COPD*, **10**, 324–332.
5. López Varela, M.V., Montes de Oca, M., Halbert, R., Muiño, A., Tálamo, C., Pérez-Padilla, R., Jardim, J.R.B., Valdivia, G., Pertuzé, J. and Menezes, A.M.B. (2013) Comorbidities and health status in individuals with and without COPD in five Latin American cities: the PLATINO study. *Arch. Bronconeumol (English Edition)*, **49**, 468–474.
6. Koskela, J., Kilpeläinen, M., Kupiainen, H., Mazur, W., Sintonen, H., Boezen, M., Lindqvist, A., Postma, D. and Laitinen, T. (2014) Co-morbidities are the key nominators of the health related quality of life in mild and moderate COPD. *BMC Pulm. Med.*, **14**, 102.
7. Frei, A., Muggensturm, P., Putcha, N., Siebeling, L., Zoller, M., Boyd, C.M., ter Riet, G. and Puhan, M.A. (2014) Five comorbidities reflected the health status in patients with chronic obstructive pulmonary disease: the newly developed COMCOLD index. *J. Clin. Epidemiol.*, **67**, 904–911.
8. Miller, J., Edwards, L.D., Agustí, A., Bakke, P., Calverley, P.M.A., Celli, B., Coxson, H.O., Crim, C., Lomas, D.A., Miller, B.E. *et al.* (2013) Comorbidity, systemic inflammation and outcomes in the ECLIPSE cohort. *Respir. Med.*, **107**, 1376–1384.
9. Centers for Disease Control and Prevention Public Health Strategic Framework for COPD Prevention. Public Health

Strategic Framework for COPD Prevention. https://www.cdc.gov/copd/resources.htm (accessed 20 March 2017).
10. Atlantis, E., Fahey, P., Cochrane, B. and Smith, S. (2013) Bidirectional associations between clinically relevant depression or anxiety and COPD: a systematic review and meta-analysis. *Chest*, **144**, 766–777.
11. Pumar, M.I., Gray, C.R., Walsh, J.R., Yang, I.A., Rolls, T.A. and Ward, D.L. (2014) Anxiety and depression—important psychological comorbidities of COPD. *J. Thorac. Dis.*, **6**, 1615–1631.
12. Pelgrim, C.E., Peterson, J.D., Gosker, H.R., Schols, A.M.W.J., van Helvoort, A., Garssen, J., Folkerts, G. and Kraneveld, A.D. (2019) Psychological co-morbidities in COPD: targeting systemic inflammation, a benefit for both? *Eur. J. Pharmacol.*, **842**, 99–110.
13. Matte, D.L., Pizzichini, M.M.M., Hoepers, A.T.C., Diaz, A.P., Karloh, M., Dias, M. and Pizzichini, E. (2016) Prevalence of depression in COPD: a systematic review and meta-analysis of controlled studies. *Respir. Med.*, **117**, 154–161.
14. Yohannes, A.M. and Alexopoulos, G.S. (2014) Depression and anxiety in patients with COPD. *Eur. Respir. Rev.*, **23**, 345–349.
15. Yohannes, A.M., Willgoss, T.G., Baldwin, R.C. and Connolly, M.J. (2010) Depression and anxiety in chronic heart failure and chronic obstructive pulmonary disease: prevalence, relevance, clinical implications and management principles. *Int. J. Geriatr. Psychiatry*, **25**, 1209–1221.
16. Montserrat-Capdevila, J., Godoy, P., Marsal, J.R., Barbé, F., Pifarré, J., Alsedà, M. and Ortega, M. (2017) Overview of the impact of depression and anxiety in chronic obstructive pulmonary disease. *Lung*, **195**, 77–85.
17. Alqahtani, J.S., Njoku, C.M., Bereznicki, B., Wimmer, B.C., Peterson, G.M., Kinsman, L., Aldabayan, Y.S., Alrajeh, A.M., Aldhahir, A.M., Mandal, S. *et al.* (2020) Risk factors for all-cause hospital readmission following exacerbation of COPD: a systematic review and meta-analysis. *Eur. Respir. Rev.*, **29**, 190166.
18. Paine, N.J., Bacon, S.L., Bourbeau, J., Tan, W.C., Lavoie, K.L., Aaron, S.D., Chapman, K.R., FitzGerald, J.M., Hernandez, P., Marciniuk, D.D. *et al.* (2019) Psychological distress is related to poor health behaviours in COPD and non-COPD patients: evidence from the CanCOLD study. *Respir. Med.*, **146**, 1–9.
19. Jang, S.M., Kim, K.U., Na, H.J., Song, S.E., Lee, S.H., Lee, H., Kim, Y.S., Lee, M.K. and Park, H.-K. (2018) Depression is a major determinant of both disease-specific and generic health-related quality of life in people with severe COPD. *Chron. Respir. Dis.*, **16**, 1479972318775422.
20. Miravitlles, M. and Ribera, A. (2017) Understanding the impact of symptoms on the burden of COPD. *Respir. Res.*, **18**, 67.
21. Montserrat-Capdevila, J., Godoy, P., Marsal, J.R., Ortega, M., Pifarré, J., Alsedà, M., Castañ, M.T. and Barbé, F. (2018) Mental disorders in chronic obstructive pulmonary diseases. *Perspect. Psychiatr. Care*, **54**, 398–404.
22. Zhou, J.J., Cho, M.H., Castaldi, P.J., Hersh, C.P., Silverman, E.K. and Laird, N.M. (2013) Heritability of chronic obstructive pulmonary disease and related phenotypes in smokers. *Am. J. Respir. Crit. Care Med.*, **188**, 941–947.
23. Kendler, K.S., Ohlsson, H., Lichtenstein, P., Sundquist, J. and Sundquist, K. (2018) The genetic epidemiology of treated major depression in Sweden. *AJP*, **175**, 1137–1144.
24. Baselmans, B.M.L., Yengo, L., van Rheenen, W. and Wray, N.R. (2021) Risk in relatives, heritability, SNP-based heritability and genetic correlations in psychiatric disorders: a review. *Biol. Psychiatry*, **89**, 11–19.
25. Fernandez-Pujals, A.M., Adams, M.J., Thomson, P., McKechanie, A.G., Blackwood, D.H.R., Smith, B.H., Dominiczak, A.F., Morris, A.D., Matthews, K., Campbell, A. *et al.* (2015) Epidemiology and heritability of major depressive disorder, stratified by age of onset, sex, and illness course in Generation Scotland: Scottish Family Health Study (GS:SFHS). *PLoS One*, **10**, e0142197.
26. Hall, R., Hall, I.P. and Sayers, I. (2019) Genetic risk factors for the development of pulmonary disease identified by genome-wide association. *Respirology*, **24**, 204–214.
27. Busch, R., Cho, M.H. and Silverman, E.K. (2017) Progress in disease progression genetics: dissecting the genetic origins of lung function decline in COPD. *Thorax*, **72**, 389–390.
28. Klimentidis, Y.C., Vazquez, A.I., de Los Campos, G., Allison, D.B., Dransfield, M.T. and Thannickal, V.J. (2013) Heritability of pulmonary function estimated from pedigree and whole-genome markers. *Front. Genet.*, **4**, 174.
29. Lu, Y., Feng, L., Feng, L., Nyunt, M.S., Yap, K.B. and Ng, T.P. (2013) Systemic inflammation, depression and obstructive pulmonary function: a population-based study. *Respir. Res.*, **14**, 1–8.
30. Al-shair, K., Kolsum, U., Dockry, R., Morris, J., Singh, D. and Vestbo, J. (2011) Biomarkers of systemic inflammation and depression and fatigue in moderate clinically stable COPD. *Respir. Res.*, **12**, 1–6.
31. Riblet, N.B., Gottlieb, D.J., Hoyt, J.E., Watts, B.V. and Shiner, B. (2020) An analysis of the relationship between chronic obstructive pulmonary disease, smoking and depression in an integrated healthcare system. *Gen. Hosp. Psychiatry*, **64**, 72–79.
32. Fluharty, M., Taylor, A.E., Grabski, M. and Munafò, M.R. (2017) The association of cigarette smoking with depression and anxiety: a systematic review. *Nicotine Tob. Res.*, **19**, 3–13.
33. Audrain-McGovern, J., Leventhal, A.M. and Strong, D.R. (2015) Chapter Eight - the role of depression in the uptake and maintenance of cigarette smoking. In De Biasi, M. (ed), *International Review of Neurobiology, Nicotine Use in Mental Illness and Neurological Disorders*. Academic Press, Waltham, MA, Vol. 124, pp. 209–243.
34. Ishii, T., Wakabayashi, R., Kurosaki, H., Gemma, A. and Kida, K. (2011) Association of serotonin transporter gene variation with smoking, chronic obstructive pulmonary disease, and its depressive symptoms. *J. Hum. Genet.*, **56**, 41–46.
35. Shrine, N., Guyatt, A.L., Erzurumluoglu, A.M., Jackson, V.E., Hobbs, B.D., Melbourne, C.A., Batini, C., Fawcett, K.A., Song, K., Sakornsakolpat, P. *et al.* (2019) New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat. Genet.*, **51**, 481–493.
36. Heinzman, J.T., Hoth, K.F., Cho, M.H., Sakornsakolpat, P., Regan, E.A., Make, B.J., Kinney, G.L., Wamboldt, F.S., Holm, K.E., Bormann, N. *et al.* (2019) GWAS and systems biology analysis of depressive symptoms among smokers from the COPDGene cohort. *J. Affect. Disord.*, **243**, 16–22.
37. Arnau-Soler, A., Macdonald-Dunlop, E., Adams, M.J., Clarke, T.-K., MacIntyre, D.J., Milburn, K., Navrady, L., Hayward, C., McIntosh, A.M. and Thomson, P.A. (2019) Genome-wide by environment interaction studies of depressive symptoms

and psychosocial stress in UK Biobank and Generation Scotland. *Transl. Psychiatry*, **9**, 1–13.
38. Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M. and Smoller, J.W. (2013) Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.*, **14**, 483–495.
39. Eid, R.S., Gobinath, A.R. and Galea, L.A.M. (2019) Sex differences in depression: insights from clinical and preclinical studies. *Prog. Neurobiol.*, **176**, 86–102.
40. Labonté, B., Engmann, O., Purushothaman, I., Menard, C., Wang, J., Tan, C., Scarpa, J.R., Moy, G., Loh, Y.-H.E., Cahill, M. *et al.* (2017) Sex-specific transcriptional signatures in human depression. *Nat. Med.*, **23**, 1102–1111.
41. Kang, H.-J., Park, Y., Yoo, K.-H., Kim, K.-T., Kim, E.-S., Kim, J.-W., Kim, S.-W., Shin, I.-S., Yoon, J.-S., Kim, J.H. *et al.* (2020) Sex differences in the genetic architecture of depression. *Sci. Rep.*, **10**, 9927.
42. Sørheim, I.-C., Johannessen, A., Gulsvik, A., Bakke, P.S., Silverman, E.K. and DeMeo, D.L. (2010) Gender differences in COPD: are women more susceptible to smoking effects than men? *Thorax*, **65**, 480–485.
43. Gan, W.Q., Man, S.P., Postma, D.S., Camp, P. and Sin, D.D. (2006) Female smokers beyond the perimenopausal period are at increased risk of chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Respir. Res.*, **7**, 52.
44. Raghavan, D. and Jain, R. (2016) Increasing awareness of sex differences in airway diseases. *Respirology*, **21**, 449–459.
45. Dashti, H.S., Daghlas, I., Lane, J.M., Huang, Y., Udler, M.S., Wang, H., Ollila, H.M., Jones, S.E., Kim, J., Wood, A.R. *et al.* (2021) Genetic determinants of daytime napping and effects on cardiometabolic health. *Nat. Commun.*, **12**, 900.
46. Mosley, J.D., Levinson, R.T., Farber-Eger, E., Edwards, T.L., Hellwege, J.N., Hung, A.M., Giri, A., Shuey, M.M., Shaffer, C.M., Shi, M. *et al.* (2020) The polygenic architecture of left ventricular mass mirrors the clinical epidemiology. *Sci. Rep.*, **10**, 7561.
47. Salem, J.-E., Shoemaker, M.B., Bastarache, L., Shaffer, C.M., Glazer, A.M., Kroncke, B., Wells, Q.S., Shi, M., Straub, P., Jarvik, G.P. *et al.* (2019) Association of thyroid function genetic predictors with atrial fibrillation: a phenome-wide association study and inverse-variance weighted average meta-analysis. *JAMA Cardiol.*, **4**, 136–143.
48. Sakornsakolpat, P., Prokopenko, D., Lamontagne, M., Reeve, N.F., Guyatt, A.L., Jackson, V.E., Shrine, N., Qiao, D., Bartz, T.M., Kim, D.K. *et al.* (2019) Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat. Genet.*, **51**, 494–505.
49. Caramori, G., Ruggeri, P., Di Stefano, A., Mumby, S., Girbino, G., Adcock, I.M. and Kirkham, P. (2018) Autoimmunity and COPD: clinical implications. *Chest*, **153**, 1424–1431.
50. Barnes, P.J. (2017) Cellular and molecular mechanisms of asthma and COPD. *Clin. Sci.*, **131**, 1541–1558.
51. Wohleb, E.S., Franklin, T., Iwata, M. and Duman, R.S. (2016) Integrating neuroimmune systems in the neurobiology of depression. *Nat. Rev. Neurosci.*, **17**, 497–511.
52. Otte, C., Gold, S.M., Penninx, B.W., Pariante, C.M., Etkin, A., Fava, M., Mohr, D.C. and Schatzberg, A.F. (2016) Major depressive disorder. *Nat. Rev. Dis. Primers.*, **2**, 1–20.
53. Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D.M., Chen, F., Datta, G., Davila-Velderrain, J., McGuire, D., Tian, C. *et al.* (2019) Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.*, **51**, 237–244.
54. Turley, P., Walters, R.K., Maghzian, O., Okbay, A., Lee, J.J., Fontana, M.A., Nguyen-Viet, T.A., Wedow, R., Zacher, M., Furlotte, N.A. *et al.* (2018) Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.*, **50**, 229–237.
55. Wootton, R.E., Richmond, R.C., Stuijfzand, B.G., Lawn, R.B., Sallis, H.M., Taylor, G.M.J., Hemani, G., Jones, H.J., Zammit, S., Davey Smith, G. *et al.* (2020) Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. *Psychol. Med.*, **50**, 2435–2443.
56. Cai, N., Revez, J.A., Adams, M.J., Andlauer, T.F.M., Breen, G., Byrne, E.M., Clarke, T.-K., Forstner, A.J., Grabe, H.J., Hamilton, S.P. *et al.* (2020) Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.*, **52**, 437–447.
57. Howard, D.M., Adams, M.J., Clarke, T.-K., Hafferty, J.D., Gibson, J., Shirali, M., Coleman, J.R.I., Hagenaars, S.P., Ward, J., Wigmore, E.M. *et al.* (2019) Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.*, **22**, 343–352.
58. Nagel, M., Jansen, P.R., Stringer, S., Watanabe, K., de Leeuw, C.A., Bryois, J., Savage, J.E., Hammerschlag, A.R., Skene, N.G., Muñoz-Manchado, A.B. *et al.* (2018) Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat. Genet.*, **50**, 920–927.
59. Goodwin, R.D., Lavoie, K.L., Lemeshow, A.R., Jenkins, E., Brown, E.S. and Fedoronko, D.A. (2012) Depression, anxiety, and COPD: the unexamined role of nicotine dependence. *Nicotine Tob. Res.*, **14**, 176–183.
60. Lou, P., Chen, P., Zhang, P., Yu, J., Wang, Y., Chen, N., Zhang, L., Wu, H. and Zhao, J. (2014) Effects of smoking, depression, and anxiety on mortality in COPD patients: a prospective study. *Respir. Care*, **59**, 54–61.
61. Mathew, A.R., Yount, S.E., Kalhan, R. and Hitsman, B. (2018) Psychological functioning in patients with chronic obstructive pulmonary disease: a preliminary study of relations with smoking status and disease impact. *Nicotine Tob. Res.*, **21**, 686–690.
62. Callaghan, R.C., Veldhuizen, S., Jeysingh, T., Orlan, C., Graham, C., Kakouris, G., Remington, G. and Gatley, J. (2014) Patterns of tobacco-related mortality among individuals diagnosed with schizophrenia, bipolar disorder, or depression. *J. Psychiatr. Res.*, **48**, 102–110.
63. Tidey, J.W. and Miller, M.E. (2015) Smoking cessation and reduction in people with chronic mental illness. *BMJ*, **351**, h4065
64. Tam, J., Warner, K.E. and Meza, R. (2016) Smoking and the reduced life expectancy of individuals with serious mental illness. *Am. J. Prev. Med.*, **51**, 958–966.
65. Wray, N.R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E.M., Abdellaoui, A., Adams, M.J., Agerbo, E., Air, T.M., Andlauer, T.M.F. *et al.* (2018) Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.*, **50**, 668–681.
66. Howard, D.M., Adams, M.J., Shirali, M., Clarke, T.-K., Marioni, R.E., Davies, G., Coleman, J.R.I., Alloza, C., Shen, X., Barbu, M.C. *et al.* (2018) Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat. Commun.*, **9**, 1470.

67. Martin, A.R., Gignoux, C.R., Walters, R.K., Wojcik, G.L., Neale, B.M., Gravel, S., Daly, M.J., Bustamante, C.D. and Kenny, E.E. (2017) Human demographic history impacts genetic risk prediction across diverse populations. *Am. J. Hum. Genet.*, **100**, 635–649.
68. Duncan, L., Shen, H., Gelaye, B., Meijsen, J., Ressler, K., Feldman, M., Peterson, R. and Domingue, B. (2019) Analysis of polygenic risk score usage and performance in diverse human populations. *Nat. Commun.*, **10**, 3328.
69. Lutz, S.M., Cho, M.H., Young, K., Hersh, C.P., Castaldi, P.J., McDonald, M.-L., Regan, E., Mattheisen, M., DeMeo, D.L., Parker, M. *et al.* (2015) A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. *BMC Genet.*, **16**, 138.
70. Wyss, A.B., Sofer, T., Lee, M.K., Terzikhan, N., Nguyen, J.N., Lahousse, L., Latourelle, J.C., Smith, A.V., Bartz, T.M., Feitosa, M.F. *et al.* (2018) Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function. *Nat. Commun.*, **9**, 2976.
71. Wei, W.-Q. and Denny, J.C. (2015) Extracting research-quality phenotypes from electronic health records to support precision medicine. *Genome Med.*, **7**, 41.
72. Chan, K.S., Fowles, J.B. and Weiner, J.P. (2010) Review: electronic health records and the reliability and validity of quality measures: a review of the literature. *Med. Care Res. Rev.*, **67**, 503–527.
73. Hersh, W.R., Weiner, M.G., Embi, P.J., Logan, J.R., Payne, P.R.O., Bernstam, E.V., Lehmann, H.P., Hripcsak, G., Hartzog, T.H., Cimino, J.J. *et al.* (2013) Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med. Care*, **51**, S30–S37.
74. Weiskopf, N.G. and Weng, C. (2013) Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J. Am. Med. Inform. Assoc.*, **20**, 144–151.
75. Hripcsak, G. and Albers, D.J. (2013) Next-generation phenotyping of electronic health records. *J. Am. Med. Inform. Assoc.*, **20**, 117–121.
76. Wei, W.-Q., Bastarache, L.A., Carroll, R.J., Marlo, J.E., Osterman, T.J., Gamazon, E.R., Cox, N.J., Roden, D.M. and Denny, J.C. (2017) Evaluating phecodes, clinical classification software, and ICD-9-CM codes for phenome-wide association studies in the electronic health record. *PLoS One*, **12**, e0175508.
77. Beesley, L.J., Fritsche, L.G. and Mukherjee, B. (2020) An analytic framework for exploring sampling and observation process biases in genome and phenome-wide association studies using electronic health records. *Stat. Med.*, **39**, 1965–1979.
78. Funk, M.J. and Landi, S.N. (2014) Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. *Curr Epidemiol Rep*, **1**, 175–185.
79. Polubriaginof, F., Salmasian, H., Albert, D.A. and Vawdrey, D.K. (2018) Challenges with collecting smoking status in electronic health records. *AMIA Annu. Symp. Proc.*, **2017**, 1392–1400.
80. Szatkowski, L., Lewis, S., McNeill, A. and Coleman, T. (2010) Is smoking status routinely recorded when patients register with a new GP? *Fam. Pract.*, **27**, 673–675.
81. Marston, L., Carpenter, J.R., Walters, K.R., Morris, R.W., Nazareth, I., White, I.R. and Petersen, I. (2014) Smoker, ex-smoker or non-smoker? The validity of routinely recorded smoking status in UK primary care: a cross-sectional study. *BMJ Open*, **4**, e004958.
82. Wu, C.-Y., Chang, C.-K., Robson, D., Jackson, R., Chen, S.-J., Hayes, R.D. and Stewart, R. (2013) Evaluation of smoking status identification using electronic health records and open-text information in a large mental health case register. *PLoS One*, **8**, e74262.
83. Roden, D., Pulley, J., Basford, M., Bernard, G., Clayton, E., Balser, J.R. and Masys, D.R. (2008) Development of a large-scale de-identified DNA Biobank to enable personalized medicine. *Clin. Pharmacol. Ther.*, **84**, 362–369.
84. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
85. Patterson, N., Price, A.L. and Reich, D. (2006) Population structure and eigenanalysis. *PLoS Genet.*, **2**, e190.
86. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M. *et al.* (2016) Next-generation genotype imputation service and methods. *Nat. Genet.*, **48**, 1284–1287.
87. Purcell, S. and Chang, C. PLINK [v1.9]. www.cog-genomics.org/plink/1.9/.
88. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. and Lee, J.J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, **4**, 7.
89. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L. and Neale, B.M. (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.*, **47**, 291–295.
90. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.-R., Duncan, L., Perry, J.R.B., Patterson, N., Robinson, E.B. *et al.* (2015) An atlas of genetic correlations across human diseases and traits. *Nat. Genet.*, **47**, 1236–1241.
91. Shi, H., Mancuso, N., Spendlove, S. and Pasaniuc, B. (2017) Local genetic correlation gives insights into the shared genetic architecture of complex traits. *Am. J. Hum. Genet.*, **101**, 737–751.
92. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C.A. and Smoller, J.W. (2019) Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.*, **10**, 1776.
93. Bush, W.S., Oetjens, M.T. and Crawford, D.C. (2016) Unravelling the human genome-phenome relationship using phenome-wide association studies. *Nat. Rev. Genet.*, **17**, 129–145.
94. Wu, P., Gifford, A., Meng, X., Li, X., Campbell, H., Varley, T., Zhao, J., Carroll, R., Bastarache, L., Denny, J.C. *et al.* (2019) Mapping ICD-10 and ICD-10-CM codes to phecodes: workflow development and initial evaluation. *JMIR Med. Inform.*, **7**, e14325.
95. Denny, J.C., Bastarache, L. and Roden, D.M. (2016) Phenome-wide association studies as a tool to advance precision medicine. *Annu. Rev. Genomics Hum. Genet.*, **17**, 353–373.
96. Denny, J.C., Bastarache, L., Ritchie, M.D., Carroll, R.J., Zink, R., Mosley, J.D., Field, J.R., Pulley, J.M., Ramirez, A.H., Bowton, E. *et al.* (2013) Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat. Biotechnol.*, **31**, 1102–1110.
97. Fritsche, L.G., Beesley, L.J., VandeHaar, P., Peng, R.B., Salvatore, M., Zawistowski, M., Taliun, S.A.G., Das, S., LeFaive, J., Kaleba, E.O. *et al.* (2019) Exploring various polygenic risk scores for skin cancer in the phenomes of the Michigan genomics initiative and the UK Biobank with a visual catalog: PRSWeb. *PLoS Genet.*, **15**, e1008202.

98. Fritsche, L.G., Gruber, S.B., Wu, Z., Schmidt, E.M., Zawistowski, M., Moser, S.E., Blanc, V.M., Brummett, C.M., Kheterpal, S., Abecasis, G.R. *et al.* (2018) Association of polygenic risk scores for multiple cancers in a phenome-wide study: results from the Michigan Genomics Initiative. *Am. J. Hum. Genet.*, **102**, 1048–1061.
99. Fritsche, L.G., Patil, S., Beesley, L.J., VandeHaar, P., Salvatore, M., Ma, Y., Peng, R.B., Taliun, D., Zhou, X. and Mukherjee, B. (2020) Cancer PRSweb: an online repository with polygenic risk scores for major cancer traits and their evaluation in two independent biobanks. *Am. J. Hum. Genet.*, **107**, 815–836.
100. Joo, Y.Y., Actkins, K., Pacheco, J.A., Basile, A.O., Carroll, R., Crosslin, D.R., Day, F., Denny, J.C., Velez Edwards, D.R., Hakonarson, H. *et al.* (2020) A polygenic and phenotypic risk prediction for polycystic ovary syndrome evaluated by phenome-wide association studies. *J. Clin. Endocrinol. Metab.*, **105**, 1918–1936.
101. Kember, R.L., Merikangas, A.K., Verma, S.S., Verma, A., Judy, R., Center, R.G., Damrauer, S.M., Ritchie, M.D., Rader, D.J. and Bućan, M. (2021) Polygenic risk of psychiatric disorders exhibits cross-trait associations in electronic health record data. *Biol. Psychiatry*, **89**, 236–245.
102. Krapohl, E., Euesden, J., Zabaneh, D., Pingault, J.-B., Rimfeld, K., von Stumm, S., Dale, P.S., Breen, G., O'Reilly, P.F. and Plomin, R. (2016) Phenome-wide analysis of genome-wide polygenic scores. *Mol. Psychiatry*, **21**, 1188–1193.
103. Leppert, B., Millard, L.A.C., Riglin, L., Davey Smith, G., Thapar, A., Tilling, K., Walton, E. and Stergiakouli, E. (2020) A cross-disorder PRS-pheWAS of 5 major psychiatric disorders in UK Biobank. *PLoS Genet.*, **16**, e1008185.
104. Li, R., Chen, Y., Ritchie, M.D. and Moore, J.H. (2020) Electronic health records and polygenic risk scores for predicting disease risk. *Nat. Rev. Genet.*, **21**, 493–502.
105. Robinson, J.R., Wei, W.-Q., Roden, D.M. and Denny, J.C. (2018) Defining phenotypes from clinical data to drive genomic research. *Ann. Rev. Biomed. Data Sci.*, **1**, 69–92.
106. Roden, D.M. (2017) Phenome-wide association studies: a new method for functional genomics in humans. *J. Physiol.*, **595**, 4109–4115.
107. Shen, X., Howard, D.M., Adams, M.J., Hill, W.D., Clarke, T.-K., Deary, I.J., Whalley, H.C. and McIntosh, A.M. (2020) A phenome-wide association and Mendelian randomisation study of polygenic risk for depression in UK Biobank. *Nat. Commun.*, **11**, 2301.
108. Zheutlin, A.B., Dennis, J., Karlsson Linnér, R., Moscati, A., Restrepo, N., Straub, P., Ruderfer, D., Castro, V.M., Chen, C.-Y., Ge, T. *et al.* (2019) Penetrance and pleiotropy of polygenic risk scores for schizophrenia in 106,160 patients across four health care systems. *Am. J. Psychiatry*, **176**, 846–855.
109. Carroll, R.J., Bastarache, L. and Denny, J.C. (2014) R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinformatics*, **30**, 2375–2376.
110. Zhu, Z., Zheng, Z., Zhang, F., Wu, Y., Trzaskowski, M., Maier, R., Robinson, M.R., McGrath, J.J., Visscher, P.M., Wray, N.R. *et al.* (2018) Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.*, **9**, 224.
111. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E. *et al.* (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.*, **47**, D1005–D1012.
112. Leslie, R., O'Donnell, C.J. and Johnson, A.D. (2014) GRASP: analysis of genotype-phenotype results from 1390 genome-wide association studies and corresponding open access database. *Bioinformatics*, **30**, i185–i194.