Lung function and cardiovascular disease: a two-sample Mendelian randomisation study

Daniel H. Higbee[1,2], Raquel Granell[1], Eleanor Sanderson[1], George Davey Smith[1] and James W. Dodd[1,2]

[1]MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, UK. [2]Academic Respiratory Unit, University of Bristol, Southmead Hospital, Bristol, UK.

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

Shareable abstract (@ERSpublications)
This two-sample multivariable Mendelian randomisation study provides strong evidence that FVC (but not FEV_1 or FEV_1/FVC <0.7) is causally associated with coronary artery disease https://bit.ly/3t05kWJ

Cite this article as: Higbee DH, Granell R, Sanderson E, *et al*. Lung function and cardiovascular disease: a two-sample Mendelian randomisation study. *Eur Respir J* 2021; 58: 2003196 [DOI: 10.1183/13993003.03196-2020].

This article has supplementary material available from erj.ersjournals.com

Received: 18 Aug 2020
Accepted: 21 Jan 2021

Abstract

Background Observational studies suggest an association between reduced lung function and risk of coronary artery disease and ischaemic stroke, independent of shared cardiovascular risk factors such as cigarette smoking. We use the latest genetic epidemiological methods to determine whether impaired lung function is causally associated with an increased risk of cardiovascular disease.

Methods and findings Mendelian randomisation uses genetic variants as instrumental variables to investigate causation. Preliminary analysis used two-sample Mendelian randomisation with lung function single nucleotide polymorphisms. To avoid collider bias, the main analysis used single nucleotide polymorphisms for lung function identified from UKBiobank in a multivariable Mendelian randomisation model conditioning for height, body mass index and smoking.

Multivariable Mendelian randomisation shows strong evidence that reduced forced vital capacity (FVC) causes increased risk of coronary artery disease (OR 1.32, 95% CI 1.19–1.46 per standard deviation). Reduced forced expiratory volume in 1 s (FEV_1) is unlikely to cause increased risk of coronary artery disease, as evidence of its effect becomes weak after conditioning for height (OR 1.08, 95% CI 0.89–1.30). There is weak evidence that reduced lung function increases risk of ischaemic stroke.

Conclusion There is strong evidence that reduced FVC is independently and causally associated with coronary artery disease. Although the mechanism remains unclear, FVC could be taken into consideration when assessing cardiovascular risk and considered a potential target for reducing cardiovascular events. FEV_1 and airflow obstruction do not appear to cause increased cardiovascular events; confounding and collider bias may explain previous findings of a causal association.

Introduction

Multimorbidity, the co-existence of multiple diseases in an individual, is associated with poor quality of life, mortality and polypharmacy [1]. Impaired lung function measures such as forced expiratory volume in 1 s (FEV_1) and forced vital capacity (FVC) have been found to be strongly associated with multimorbidity and are reported as independent predictors of cardiovascular disease [2]. Although research has often focused on the contribution of FEV_1 and obstructive airways disease to cardiovascular risk, FVC has been shown to be a stronger predictor of survival, and appears to add value to the Framingham risk score for prediction of mortality [3, 4]. However, it is unclear if there is a causal link between lung function and multimorbidity, or if the association is due to confounding factors such as cigarette smoking.

Observational studies have reported that COPD, decreased FEV_1, FVC and FEV_1/FVC ratio are all associated with an increased the risk of coronary artery disease (CAD) [5, 6]. However, results are inconsistent, with some studies reporting no association [7], or that the association is limited to those with

abnormally high blood pressure [8]. There is evidence suggesting that COPD and impaired lung function are associated with an increased risk of stroke [9].

Impaired lung function and associated lung diseases could have a direct detrimental effect on cardiovascular health *via* a number of biological pathways including systemic inflammation or oxidative stress [10, 11]. However, the mechanisms may vary between different lung function traits [12].

Mendelian randomisation (MR) is a method that can overcome problems of unmeasured confounding and reverse causation typical of conventional observational epidemiology [13]. MR allows causal inference through the use of genetic variants as proxies for modifiable risk factors or health outcomes [14]. MR uses genetic data, *e.g.* single nucleotide polymorphisms (SNPs) that are associated with an exposure (in this case lung function) as instrumental variables to assess the causal effect of the exposure on the outcome of interest (in this case cardiovascular disease) [15].

MR has multiple advantages; it uses genetic variants, which are randomly allocated at conception, so they can be exploited to simulate randomisation [15]. Genetic variants are not influenced by behavioural or environmental factors and are far less susceptible to bias from reverse causation. Additionally, the effects are equivalent to lifetime differences, reducing issues relating to transient fluctuations in exposures [16]. Multivariable MR (MVMR) has further advantages; it includes multiple exposures in the model, allowing estimation of the direct causal effect of each exposure on the outcome. Each exposure SNP has its effect on all exposures, *e.g.* lung function trait and height included in the MR model, allowing for conditioning. MVMR is a robust method when using two exposures that could act as a confounder, mediator or collider of the exposure–outcome relationship [17, 18]. Our objective was to determine whether impaired lung function causally increases the risk of cardiovascular disease.

Methods

Exposures: Shrine et al. preliminary analysis

We used data from the largest currently available lung function genome-wide association study (GWAS), by Shrine *et al.* [19], to undertake a preliminary two-sample MR analysis. The Shrine *et al.* GWAS reported 279 genome-wide significant SNPs ($p<5\times10^{-9}$) in a European ancestry population and was adjusted for age, age^2, height and smoking status. Full details are provided elsewhere [19].

The Shrine *et al.* [19] GWAS adjusted for covariates of lung function and cardiovascular disease, *e.g.* height and smoking; this can lead to collider bias as SNPs can be related to both the covariates, *e.g.* height and to other adverse risk factors [16]. This can lead to false-positive SNP discoveries and bias (towards null effect) in MR studies [20].

Exposures: main analysis MVMR

To avoid the collider bias we used exposure SNPs discovered in GWAS that had not been adjusted for covariates in an MVMR model. To find suitable exposure SNPs we used the UKBiobank of 502 543 individuals aged between 40 and 69 years at recruitment across the UK (www.ukbiobank.ac.uk/). Participants completed detailed health questionnaires and blood samples were taken for genotyping. Of these, 353 315 participants have "best measures" of pre-bronchodilator FEV_1 and FVC, measured as absolute values in litres. We performed a GWAS on these individuals (adjusting for sex). In addition, we performed a GWAS based on 55 907 cases of airflow obstruction (defined as FEV_1/FVC <0.70) and 297 408 controls (FEV_1/FVC $\geqslant$0.70). The SNPs discovered in this unadjusted GWAS were then used in a two-sample MVMR model conditioning with SNPs for covariates of exposure and outcome: standing height, body mass index (BMI) and current smoking. SNPs for these covariates were identified in pre-existing GWAS performed in the UKBiobank (https://gwas.mrcieu.ac.uk/). Details are provided in the supplementary material. Note that the function of genetic variants is independent of age, and adjusting for it in a two-sample MR model is not necessary or possible (as age is not genetically determined). All exposure SNPs were discovered in only European ancestry populations.

Outcomes

We used the CARDIoGRAMplusC4D GWAS based on 60 901 cases of CAD and 123 504 controls, 77% of whom were of European ancestry [21]. CAD was defined by myocardial infarction, acute coronary syndrome, chronic stable angina or coronary stenosis of >50%.

For stroke we used the MEGASTROKE GWAS based on 34 217 cases of acute ischaemic stroke and 406 111 controls, all of European ancestry [22]. There was no overlap between our exposure and outcome population samples.

Statistical analysis

Statistical analysis was done using R Studio (version 3.6.1, www.rstudio.com) with MRCIEU/TwoSampleMR and MRInstruments packages [17, 23].

F-statistics were calculated to assess exposure instruments strength [24]. Linkage disequilibrium (LD) clumping and Steiger filtering were performed [23]. Duplicate SNPs and palindromic SNPs were removed, and all SNPs were harmonised. Proxies were identified when CAD was the outcome. Details are provided in appendix 3.

Main MR analysis

Inverse variance weighting (IVW) was used for main effect estimate for both MVMR and two-sample MR analyses. This IVW is a weighted regression of SNP–outcome on SNP–exposure associations combined.

Results

Shrine et al. *preliminary analysis*

Due to collider bias, results from this analysis should be interpreted with caution. When adjusting for a covariate, the effect estimate of the SNP with lung function will be biased by the correlation between the covariate and lung function multiplied by their association with covariate. For example, if a SNP has a strong positive effect on height it would reduce the observed effect on lung function. Adjusting for a covariate in a GWAS could induce an association between SNPs associated with the covariate and the adjusted trait that is inverse to the true association between each SNP and the covariate [20]. This bias in the SNP–exposure association will feed through to any MR estimates obtained using it and could lead to bias in the MR estimates obtained, either towards or away from the null. The implications for MR estimates from covariate-adjusted GWAS are explained in detail elsewhere [25]. A directed acyclic graph and further details are provided in appendix 8 and supplementary figure E9.

All analysis showed weak evidence of an effect, variable direction of effect and wide confidence intervals. These results are reported in further detail in the supplementary material. We proceeded with MVMR as our main analysis as a more robust method able to account for collider bias.

MVMR

Using a threshold of $p<5\times10^{-8}$, after quality control and LD-clumping, the unadjusted GWAS of lung function in UKBiobank produced 360 SNPs for FEV_1, 464 SNPs for FVC and 154 SNPs for FEV_1/FVC ratio <0.70 explaining 3.6%, 4.8% and 0.9% of variance, respectively. F-statistics for FEV_1, FVC and FEV_1/FVC ratio <0.7 were 38, 40 and 36, respectively. For covariates, F-statistics for standing height, BMI and current smoking were 50, 39 and 32, respectively.

MVMR analysis: FEV_1 and FVC as exposure, CAD as outcome

Results are presented as per SD decrease in lung function trait. Analysis showed strong evidence of an increased risk of CAD per SD decrease in FVC (OR 1.32, 95% CI 1.19–1.46 per SD), as shown in table 1. This effect did not attenuate after conditioning for BMI (1.41, 1.25–1.59) or current smoking (1.32, 1.19–1.47), but was weaker after conditioning for height (1.22, 1.03–1.44).

Prior to any conditioning, there was evidence that reduced FEV_1 increases risk of CAD (OR 1.27, 95% CI 1.12–1.44 per SD). However, when conditioning for height the effect size decreases with widening of the confidence interval which cross 1.0 (1.08, 0.89–1.30) (table 1). This is probably due to the pleiotropy in the MR analysis as the unadjusted GWAS would have discovered SNPs that affected lung function *via* height. Therefore, there is limited evidence of a direct effect of FEV_1 on cardiovascular risk. Conditioning for BMI (1.26, 1.08–1.47) and current smoking (1.26, 1.10–1.44) made minimal difference to the estimated effect.

MVMR analysis: FEV_1 and FVC as exposure, ischaemic stroke as outcome

There is little evidence to suggest that reduced FEV_1 increases the risk of ischaemic stroke (OR 1.11, 95% CI 0.97–1.26 per SD) (table 1). The magnitude decreased further when conditioning for both height and BMI, although the direction remained consistent. There is evidence that a decrease in FVC increases risk of ischaemic stroke (1.23, 1.01–1.24), but the effect size and strength of evidence attenuates after conditioning for height or BMI (1.16, 0.98–1.38 and 1.05, 0.93–1.19, respectively). Results for effects of FEV_1 and FVC on CAD and ischaemic stroke after conditioning for all covariates together are presented in appendix 4.

TABLE 1 Multivariable Mendelian randomisation results of forced expiratory volume in 1 s (FEV_1) and forced vital capacity (FVC) on coronary artery disease and ischaemic stroke using UK Biobank lung function genome-wide association study

	Condition	SNPs (LF/condition)	Coronary artery disease OR (95% CI)#	SNPs (LF/condition)	Ischaemic stroke OR (95% CI)#
FEV_1	None	300/0	1.27 (1.12–1.44)	291/0	1.11 (0.97–1.26)
FEV_1	Height	194/744	1.08 (0.89–1.30)	193/741	1.01 (0.83–1.22)
FEV_1	BMI	179/645	1.26 (1.08–1.47)	185/660	1.03 (0.88–1.20)
FEV_1	Smoking	274/15	1.26 (1.10–1.44)	273/12	1.11 (0.95–1.29)
FVC	None	391/0	1.32 (1.19–1.46)	384/0	1.12 (1.01–1.24)
FVC	Height	272/726	1.22 (1.03–1.44)	273/728	1.04 (0.88–1.24)
FVC	BMI	227/599	1.41 (1.25–1.59)	227/607	1.05 (0.93–1.19)
FVC	Smoking	359/15	1.32 (1.19–1.47)	368/11	1.11 (1.00–1.23)

Data are presented as n, unless otherwise stated. SNP: single nucleotide polymorphism; LF: lung function trait; BMI: body mass index. #: per SD decrease in LF.

MVMR analysis: FEV_1/FVC ratio <0.7 as exposure, CAD and ischaemic stroke as outcomes

Steiger filtering removed 87 SNPs for FEV_1/FVC ratio <0.7 with CAD as the outcome and 96 SNPs with ischaemic stroke as the outcome. We found very little evidence of an effect of liability to airflow obstruction on cardiovascular disease, as can be seen in table 2.

Discussion

This MVMR study provides evidence that a single SD reduction in FVC causes ~20% increased risk of CAD. This finding confirms causality of previous observational associations [5, 6]. These results are unlikely to be affected by reverse causation or confounding factors due to the use of SNPs as instrumental variables. This effect was not seen in the preliminary non-MVMR analysis because of collider bias introduced to the model by covariate adjustment in the SHRINE *et al.* discovery GWAS. Our main analysis used MVMR, which is a robust tool when a secondary exposure acts as a confounder, a mediator, a pleiotropic pathway and a collider [26].

Although historically most observational studies of cardiovascular morbidity have focused on FEV_1 and COPD, we found little evidence of a causal association between FEV_1 and liability to obstructive ratio on cardiovascular disease risk. These results mirror findings that FVC is stronger predictor of overall survival than FEV_1 [3]. Our findings suggest that the observed association between low FEV_1, obstruction and increased risk of cardiovascular disease is unlikely to be causal. In healthy individuals, FEV_1 and FVC are highly correlated. Therefore, we hypothesise that the unknown underlying biological mechanism linking lung function and cardiovascular disease may be specific to FVC reduction.

Finding modifiable risk factors for CAD is important; however, the majority of therapies designed to improve lung function (such as inhaled bronchodilators) have a temporary and limited impact on FVC and so are unlikely to be sufficient to modify cardiovascular risk. Available treatments which do target decline in FVC are for specific and rare lung disease such as pulmonary fibrosis [27].

There are a number of strengths to our study; first, it utilises large numbers of instrumental variables, far more than were available in previous MR studies [28]. Second, we used a huge exposure sample population and multiple robust methods and adhered to rigorous proposed Strengthening the Reporting of Observational Studies in Epidemiology. guidelines for MR studies [29]. By using MR, we accounted for unmeasured confounding and reverse causation, problems typical of conventional observational epidemiology and establish causality by the use of randomly assigned genetic instrumental variables [13,

TABLE 2 Multivariable Mendelian randomisation results of forced expiratory volume in 1 s (FEV_1)/forced vital capacity (FVC) ratio <0.7 on coronary artery disease and ischaemic stroke using UK Biobank lung function genome-wide association study

	Condition	SNPs (LF/condition)	Coronary artery disease OR (95% CI)#	SNPs (LF/condition)	Ischaemic stroke OR (95% CI)#
FEV_1/FVC <0.7	None	50/0	1.00 (0.60–1.67)	39/0	0.96 (0.52–1.79)
FEV_1/FVC <0.7	Smoking	49/17	1.00 (0.83–1.21)	38/13	0.98 (0.82–1.16)

Data are presented as n, unless otherwise stated. SNP: single nucleotide polymorphism; LF: lung function trait. #: per SD increase in liability to ratio <0.7.

30, 31]. In addition, our study benefited from using MVMR to condition for these covariates avoiding collider bias that could have contributed to the weak evidence found in our preliminary analysis using the SHRINE *et al.* [19] GWAS. MVMR estimates the direct rather than total effect of an exposure, allowing us to show that much of the effect of FEV_1 on CAD risk was due to pleiotropic SNPs affecting FEV_1 *via* height (an established determinant of cardiovascular risk). Finally, this is the first study to use SNPs for FEV_1/FVC <0.7 ratio.

MR has assumptions and is vulnerable to certain biases if not used properly. The sensitivity analysis using plots, MR-Egger, weighted median and mode did not indicate any violation of assumptions. The use of Steiger filtering reduces the risk of reverse causality.

Limitations

Our exposure GWAS and the MEGASTROKE used only those of European heritage. The CARDIOGRAMplusC4D GWAS was 23% non-European heritage. Lung function SNPs discovered in European ancestral populations in the Shrine GWAS have been shown to have a smaller effect in non-European populations [19]. As our own UKBiobank GWASs used a high proportion of the same sample examining similar traits, it is likely that in a non-European population the effects would be smaller. We did not have access to another sample population to estimate the effects of SNPs discovered in our GWAS. As our SNPs were discovered and effects estimated in the same population, the effects could have been overestimated due to "winner's curse" phenomena [32]. There was a reduction in number of instruments available for analysis following LD-clumping, removal of duplicates and extraction from exposure and outcome GWAS. This reduces the strength of the instruments which may have reduced the power to show an effect of FEV_1 or FEV_1/FVC <0.7 ratio. In our MVMR analysis we used FEV_1/FVC <0.7 ratio as an exposure because this is a commonly used threshold of obstructive lung function. Using FEV_1/FVC ratio as a continuous trait has inherent issues in MR analysis. High FEV_1/FVC ratio is a sign of restriction and low FEV_1/FVC ratio defines airflow obstruction, both of which are pathological states that could affect cardiovascular disease, making interpretation of the continuous variable challenging. Most MR analysis assumes a linear effect, which would be violated when using FEV_1/FVC as a continuous trait. Dichotomisation of continuous traits in MR studies can make interpretation of the causal estimate less reliable, but MR can still be a valid test of the causal null hypothesis for a binary exposure [33]. An assumption of MR is that SNPs only affect the outcome *via* the exposure. To ensure that our SNPs were not affecting our outcomes *via* amount smoked we checked to see if any of our lung function SNPs are found in the 15q25 locus [34]. In the MVMR analysis for FEV_1 only one SNP (rs72736802) is from the locus, none from the FVC analysis. Therefore, we do not think this will affect our results. Lung function is a complex trait and SNPs affect lung function *via* differing pathological processes [19]. The differing processes may vary in their impact on the risk of comorbidities, perhaps reflected in the assessments of heterogeneity. It is possible our study was limited by the number of ischaemic stroke cases in the outcome population. If there is a causal effect of lung function on ischaemic stroke, it is likely to only occur with large changes in lung function as seen with CAD.

Implications

There are several important implications of our findings, first is that it is FVC not obstructive lung function that is causally associated with CAD. This suggests that we should focus our attention on understanding the mechanisms by which FVC causes CAD. Second, given there are limited FVC-specific therapies, future interventions to improve CAD outcomes through modifying FVC are most likely to be achieved through environmental/behavioural public health interventions designed to achieve optimal lung development and preventing lung function decline. Third, FVC is a widely and routinely collected clinical measure (spirometry); this study supports the call for FVC measurements to be evaluated as part of cardiovascular prognostication/secondary prevention risk assessments.

It remains uncertain whether lung function has a causal effect on the risk of ischaemic stroke. Our MVMR models show very little weak evidence that reduced lung function increases the risk of ischaemic stroke. Larger outcome sample sizes may become available as genetic consortia grow which could provide more conclusive results. Future studies are needed to determine the mechanism by which FVC causes increased CAD.

Conclusions

There is strong evidence that reduced FVC is independently and causally associated with CAD. Although the mechanism remains unclear, FVC may play an important contribution to the assessment of cardiovascular risk. Further studies are needed to test whether interventions to improve or maintain FVC may also modify cardiovascular risk. FEV_1 and obstructive lung function do not appear to cause increased

cardiovascular events; confounding and collider bias may explain previous observational and MR findings of a causal association.

Acknowledgement: This project was undertaken using UK Biobank, application ID 55521. We would like to thank participants of all the consortia used including UK BioBank, MEGASTROKE and CARDIOGRAMplusC4D.

Author contributions: All authors contributed to analysis, interpretation and final manuscript preparation; J.W. Dodd and D.H. Higbee contributed to study concept, rationale, and initial manuscript drafts.

Conflict of interest: None declared.

Support statement: This work was supported by the Medical Research Council and the University of Bristol (MC_UU_00011/1); MRC CARP Fellowship. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 National Institute for Health and Care Excellence (NICE). Multimorbidity: Clinical Assessment and Management. 2016. www.nice.org.uk/ng56
2 Agustí 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.
3 Burney PGJ, Hooper R. Forced vital capacity, airway obstruction and survival in a general population sample from the USA. *Thorax* 2011; 66: 49–54.
4 Lee HM, Le H, Lee BT, *et al.* Forced vital capacity paired with Framingham risk score for prediction of all-cause mortality. *Eur Respir J* 2010; 36: 1002–1006.
5 Schroeder EB, Welch VL, Couper D, *et al.* Lung function and incident coronary heart disease: the Atherosclerosis Risk in Communities Study. *Am J Epidemiol* 2003; 158: 1171–1181.
6 Kim JJ, Kim DB, Jang SW, *et al.* Relationship between airflow obstruction and coronary atherosclerosis in asymptomatic individuals: evaluation by coronary CT angiography. *Int J Cardiovasc Imaging* 2018; 34: 641–648.
7 Cuttica MJ, Colangelo LA, Dransfield MT, *et al.* Lung function in young adults and risk of cardiovascular events over 29 years: the CARDIA Study. *J Am Heart Assoc* 2018; 7: e010672.
8 Engström G, Hedblad B, Valind S, *et al.* Increased incidence of myocardial infarction and stroke in hypertensive men with reduced lung function. *J Hypertens* 2001; 19: 295–301.
9 Hozawa A, Billings JL, Shahar E, *et al.* Lung function and ischemic stroke incidence: the Atherosclerosis Risk in Communities study. *Chest* 2006; 130: 1642–1649.
10 Corlateanu A, Covantev S, Mathioudakis AG, *et al.* Chronic obstructive pulmonary disease and stroke. *COPD* 2018; 15: 405–413.
11 Maclay JD, MacNee W. Cardiovascular disease in COPD: mechanisms. *Chest* 2013; 143: 798–807.
12 Ramalho SHR, Shah AM. Lung function and cardiovascular disease: a link. *Trends Cardiovasc Med* 2021; 31: 93–98.
13 Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; 32: 1–22.
14 Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014; 23: R89–R98.
15 Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* 2017; 26: 2333–2355.
16 Davey Smith G, Paternoster L, Relton C. When will Mendelian randomization become relevant for clinical practice and public health? *JAMA* 2017; 317: 589–591.
17 Sanderson E, Davey Smith G, Windmeijer F, *et al.* An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol* 2019; 48: 713–727.
18 Marouli E, Del Greco MF, Astley CM, *et al.* Mendelian randomisation analyses find pulmonary factors mediate the effect of height on coronary artery disease. *Commun Biol* 2019; 2: 119.
19 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.
20 Aschard H, Vilhjálmsson BJ, Joshi AD, *et al.* Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am J Hum Genet* 2015; 96: 329–339.
21 Nikpay M, Goel A, Won H-H, *et al.* A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015; 47: 1121–1130.
22 Malik R, Chauhan G, Traylor M, *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* 2018; 50: 524–537.

23 Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet* 2017; 13: e1007081.

24 Burgess S, Thompson SG, CRP CHD Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011; 40: 755–764.

25 Hartwig FP, Tilling K, Davey Smith G, *et al.* Bias in two-sample Mendelian randomization by using covariable-adjusted summary associations. *Int J Epidemiol* 2021; dyaa266.

26 Sanderson E, Davey Smith G, Windmeijer F, *et al.* An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol* 2019; 19: 713–727.

27 Bonella F, Stowasser S, Wollin L. Idiopathic pulmonary fibrosis: current treatment options and critical appraisal of nintedanib. *Drug Des Devel Ther* 2015; 9: 6407–6419.

28 Au Yeung SL, Borges MC, Lawlor DA. Association of genetic instrumental variables for lung function on coronary artery disease risk: a 2-sample Mendelian randomization study. *Circ Genom Precis Med* 2018; 11: e001952.

29 Davey Smith G, Davies NM, Dimou N, *et al.* STROBE-MR: guidelines for strengthening the reporting of Mendelian randomization studies. *PeerJ Preprints* 2019; preprint [https://doi.org/10.7287/peerj.preprints.27857v1]

30 Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* 2005; 330: 1076–1079.

31 Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018; 362: k601.

32 Kraft P. Curses – winner's and otherwise – in genetic epidemiology. *Epidemiology* 2008; 19: 649–651.

33 Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol* 2018; 33: 947–952.

34 The Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010; 42: 441–447.