

The impact of lung function on cardiovascular diseases and cardiovascular risk factors: a two sample bi-directional Mendelian randomization study

Au Yeung SL^{1*}, Borges MC^{2,3}, Lawlor DA^{2,3,4}, Schooling CM^{1,5}

¹ School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

² MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

³ Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

⁴ National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and University of Bristol, UK

⁵ School of Public Health and Health Policy, City University of New York, New York, USA

*Corresponding author

Au Yeung Shiu Lun

School of Public Health, LKS Faculty of Medicine, The University of Hong Kong
1/F, Patrick Manson Building, 7 Sassoon Road, Hong Kong SAR, China

Phone: 852-3917 6740

Fax: 852-3520 1945

Email: ayslryan@hku.hk

Abstract: 338

Main text: 3,717

Tables: 1

Figures: 3

Supplemental material: 2

Running title: Lung function and CVD using MR

Keywords: Atrial fibrillation; Coronary artery disease; Heart failure; Lung function;

Mendelian randomization; Stroke

Abstract

Background: Observational studies suggested lung function is inversely associated with cardiovascular disease (CVD) although these studies could be susceptible to residual confounding. We conducted a 2 sample Mendelian randomization study using summary statistics from genome wide association studies (GWAS) to clarify the role of lung function in CVD and its risk factors, and conversely the role of CVD in lung function.

Methods: We obtained genetic instruments for forced expiratory volume in 1 second (FEV₁) and forced vital capacity (FVC) from publicly available UK Biobank summary statistics (n=421,986). We applied these genetic instruments for FEV₁ (260) and FVC (320) to publicly available GWAS summary statistics for coronary artery disease (CAD) (n=184,305), stroke and its subtypes (n=446,696), atrial fibrillation (n=1,030,836), and heart failure (n=977,320) and cardiovascular risk factors. Inverse variance weighting was used to assess the impact of lung function on these outcomes. Sensitivity analyses included MR-Egger, weighted median, MR-PRESSO, and multivariable Mendelian randomization. We also conducted bi-directional Mendelian randomization to assess whether CVD affects lung function.

Results: FEV₁ and FVC were inversely associated with CAD (odds ratio (OR) per standard deviation (SD) increase, 0.72 (95% confidence interval (CI) 0.63 to 0.82) and 0.70 (95%CI 0.62 to 0.78)), overall stroke (0.87 (95%CI 0.77 to 0.97), 0.90 (0.82 to 1.00)), ischemic stroke (0.87 (95%CI 0.77 to 0.99), 0.90 (95%CI 0.80 to 1.00)), small vessel stroke (0.78, (95%CI 0.61 to 1.00), 0.74 (95%CI 0.61 to 0.92)), and large artery stroke (0.69 (95%CI 0.54 to 0.89), 0.72 (95%CI 0.57 to 0.91)). FEV₁ and FVC were inversely associated with type 2 diabetes (0.75 (95%CI 0.62 to 0.90), 0.67 (95%CI 0.58 to 0.79)) and systolic blood pressure. Sensitivity analyses produced similar direction for most outcomes although the magnitude sometimes differed. Adjusting for height attenuated results for CAD (e.g. OR for 1SD FEV₁ 0.95 (0.76 to 1.20), but this may reflect weak instrument bias. This adjustment did not attenuate effects for stroke or type 2 diabetes. No strong evidence was observed for CVD affecting lung function.

Conclusion: Higher lung function likely protect against CAD and stroke.

Introduction

Poorer lung function, indicated by reduced forced expiratory volume in 1 second (FEV_1) and forced vital capacity (FVC), is associated with cardiovascular disease (CVD).¹⁻⁵ However, these associations could be confounded by lifestyle factors, such as physical activity and smoking, and anthropometric characteristics, such as height, which are difficult to account for in observational studies.

Mendelian randomization is potentially less vulnerable to residual confounding than observational studies because it utilizes genetic variants related to exposures which are randomly allocated during conception.⁶ A previous two-sample Mendelian randomization study suggested higher FEV_1 was causally related to lower risk of coronary artery disease (CAD), whilst the relation for FVC was less clear.⁷ However, that study used a relatively small number of genetic instruments (i.e. 16 instruments for FEV_1 (variance explained (R^2): 0.8%) and 10 instruments for FVC (R^2 : 0.4%)), did not consider other important CVDs, such as stroke and heart failure, and did not explore the possibility of bidirectional effects (i.e. CVDs also having an effect on lung function).⁷ Furthermore, the genetic instruments used in that study had been adjusted for height and smoking and such adjustment can sometimes introduce bias in Mendelian randomization results.⁸ A more recent Mendelian randomization study, using UK Biobank and GWAS consortia data, concluded that most of the effect of height on CAD risk was mediated by lung function.⁹ However, that study did not examine the impact of lung function in detail (e.g. with detailed sensitivity analyses), its impact on other CVD outcomes, or possible reverse causation.

The aim of this study was to determine the causal effect of FEV₁ and FVC on a wide range of cardiovascular diseases using 2 sample Mendelian randomization with summary statistics.¹⁰ Our study adds to previous Mendelian randomization studies by including more genetic instruments and more CVD outcomes and risk factors. We also explored whether genetic predisposition to CVDs might cause variation in lung function.

Methods

Study Design

We used two sample summary data Mendelian randomization to assess the effect of FEV₁ (per standard deviation (SD)) and FVC (per SD) on multiple CVD outcomes: (i.e. CAD, stroke and its subtypes, heart failure, atrial fibrillation) and CVD risk factors (i.e. systolic and diastolic blood pressure, high density and low density lipoprotein cholesterol, triglycerides, type 2 diabetes, fasting glucose, glycated hemoglobin and insulin).¹⁰ Table 1 summarizes the study design including the sources of GWAS summary data for exposure (FEV₁ and FVC – Sample 1) and each CVD outcome and risk factors (Sample 2). Given Mendelian randomization has stringent assumptions (i.e. relevance, independence and exclusion restriction), we also assessed whether the genetic instruments affected key confounders for CVD (education, body mass index (BMI), smoking, alcohol consumption, and height) to check exclusion restriction.¹¹

Ethics approval

In this study we have used publicly-available GWAS results from relevant publications and database (<https://gwas.mrcieu.ac.uk/>).¹²⁻²⁵ No individual participant

data were collected or used. Details of ethical approval and participant consent for each of the studies that contributed to the GWAS can be found in the original publications.

Genetic determinants of FEV₁ and FVC (Sample 1)

Genetic determinants of FEV₁ (in standard deviation (SD), GWAS ID: ukb-b-19657) and FVC (SD, GWAS ID: ukb-b-7953) were extracted from summary genome-wide association study results in the UK Biobank, available in the IEU GWAS database (<https://gwas.mrcieu.ac.uk/>).^{16 17} In brief, the UK Biobank is a large prospective cohort study where more than 500,000 participants were recruited in the United Kingdom (Great Britain only) from 2006 to 2010 (mean age: 56.5; 54% female). Pre-bronchodilation lung function testing was performed by trained healthcare staff using a Viitalograph Pneumotrac 6800 spirometer (Maids Moreton, UK). Genotyping was done using the Affymetrix UK BiLEVE Axiom array (~50,000 participants) and Affymetrix UK Biobank Axiom array (~450,000 participants). The summary statistics were generated using a linear mixed model to account for relatedness and population stratification, and adjusted for sex and genotyping array. The analyses were restricted to 421,986 participants of European descent. Quality controls included exclusion based on imputation quality (specific INFO scores based on minor allele frequency (MAF)) and MAF (<=1%) (Supplemental Table 1).

Genetic associations with the outcomes (Sample 2)

Complete summary GWAS results for CAD (CARDIoGRAMplusC4D 1000 Genome based GWAS),¹² stroke and its subtypes (MEGASTROKE consortium),¹³ atrial fibrillation,¹⁴ heart failure (HERMES consortium),¹⁵ systolic and diastolic blood

pressure (UK Biobank associations obtained from IEU GWAS database),^{16 17} LDL, HDL cholesterol and triglycerides (GLGC),¹⁸ type 2 diabetes (DIAGRAM),²³ fasting glucose, insulin, and glycated hemoglobin (MAGIC),^{21 22} were obtained from publicly available online GWAS summary data repositories, with majority of them retrieved via MR Base.¹⁷ Table 1 summarizes the numbers (including the cases and controls where relevant) included in these GWAS, population (including ethnicity) and the sample size in the GWAS. Additional details, including quality control, imputation methods, and any covariates adjusted for in each GWAS and how the outcomes were defined are provided in Supplemental Table 1.

Selecting genetic instruments from the exposure GWAS, identifying these in outcome GWAS and harmonizing instruments across studies

We identified genetic instruments for lung function at genome wide significant p value ($<5 \times 10^{-8}$), and we excluded instruments which were in high linkage disequilibrium (LD) with other instruments ($r^2 < 0.001$). We searched for the genetic instruments in the outcome datasets. For genetic instruments not available for an outcome, a proxy instrument in high LD with the original instrument ($r^2 > 0.8$) was identified via MR-Base based on 1000 Genomes catalog (CEU reference population).¹⁷ No proxy instruments were identified for outcomes not available in MR-Base. We aligned each genetic association for exposure and outcome on the same effect allele. We used effect allele frequency to ensure palindromic genetic instruments were aligned properly where that was possible.

Statistical analysis

We estimated the r^2 of each genetic instrument and summed them up to compute the overall r^2 and F statistics using the sample size which the instruments for lung function was derived from (n=421,986). Higher r^2 and F statistic values suggest lower risk of weak instrument bias.^{26 27} For our main Mendelian randomization analyses, we used inverse variance weighting (IVW) with multiplicative random effects to obtain the causal effect of FEV₁ and FVC on CVD outcomes and their risk factors. IVW assumes there is no unbalanced horizontal pleiotropy.²⁸ To assess the presence of potentially invalid instruments, we checked for evidence of heterogeneity across instrument-specific Wald ratios (i.e ratio of instrument-outcome association to instrument-exposure association) using Cochran Q test. Heterogeneity may result from horizontal pleiotropy.

Additional analyses to explore potential violation of Mendelian randomization assumptions

MR assumes that genetic instruments do not affect confounders of the exposure-outcome association. We assessed the plausibility of this assumption by exploring whether the FEV₁ and FVC genetic instruments were related to key confounders defined as plausibly affecting lung function and CVDs (body mass index, height, socioeconomic position, alcohol and smoking).¹¹ This was done using publicly available summary genome wide association study data for: body mass index,²⁰ education (as a measure of socioeconomic position),²⁴ number of drinks of alcohol per week and number of cigarette smoked per day,¹⁹ and height.²⁵ Specifically, we meta analyzed the estimate of each SNP-outcome association (per lung function increasing allele) using standard inverse variance weighting method with additive random effects. For confounders which are associated with the genetic instruments,

we then assessed whether these associations were a sign of vertical pleiotropy (causal role of lung function on confounders) or horizontal pleiotropy (causal role of confounders on lung function) using bi-directional Mendelian randomization.¹¹ We used multivariable Mendelian randomization to control for potential horizontal pleiotropy by including estimates from the relevant confounders which were strongly associated with genetic instruments for lung function.^{11 29} We also approximated the conditional F statistics to evaluate potential weak instrument bias,³⁰ presented Cochrane Q test to evaluate heterogeneity, and also provided estimates for multivariable MR-Egger to assess robustness of findings due to pleiotropy.³¹

To assess unbalanced horizontal pleiotropy bias we undertook the following sensitivity analyses: MR-Egger,³² weighted median analyses,³³ and MR-PRESSO,³⁴ that are more robust than IVW to the assumption that there is no effect of the genetic instrument on outcome other than through the exposures of interest (i.e. no horizontal pleiotropy paths).²⁸ Like IVW, MR-Egger and MR-PRESSO require that the InSIDE ('instrument strength is independent of direct effect') assumption is not violated. Specifically, InSIDE assumes there is no correlation between the strength of the instrument (association of FEV₁ or FVC on their respective genetic instruments) and the strength of any association of the genetic instruments on the outcome not via the exposures of interest (e.g. any horizontal pleiotropy paths). The weighted median analyses assumed at least 50% of the weight of the genetic instruments is via the exposure of interest. Further details of these methods and their additional (and differing) assumptions are available in the supplemental material.

Bi-directional Mendelian randomization

To assess potential reverse causation, i.e. predisposition to CVDs affecting variation in FEV₁ or FVC, we also conducted a Mendelian randomization using genetic predictors of risk of CAD, overall stroke, heart failure, and atrial fibrillation as instruments.¹²⁻¹⁵ We used IVW to explore evidence of effects of genetic predispositions to these CVDs on FEV₁ or FVC.

All analyses were performed using R Version 3.5.2 (R Development Core Team, Vienna, Austria) using R packages ("TwoSampleMR"), and ("MRPRESSO").^{17 34}

Results

We included up to 260 SNPs for FEV₁, which explained 3.5% of its variance and had an overall F statistic of 58.8. For FVC, we included up to 320 SNPs, which explained 4.8% of its variance and had an overall F statistic of 66.4. Supplemental Tables 2-3 shows the summary statistics of the lung function instruments used in this study.

Figure 1 shows the IVW MR causal effect estimates of FEV₁ and FVC on the risk of CVD and type 2 diabetes. Higher FEV₁ was associated with lower risk of CAD, overall stroke and some subtypes (i.e. ischemic stroke, small vessel stroke, large artery stroke), and type 2 diabetes, but increased risk of atrial fibrillation. FVC showed similar relation with these outcomes. There was no strong evidence that either FEV₁ or FVC influenced heart failure (both point estimates were close to the null). Figure 2 shows the IVW MR causal effect estimates of FEV₁ and FVC on cardiovascular risk factors using IVW. Higher FEV₁ was mainly associated with lower systolic blood pressure and triglycerides. FVC showed similar relations with these

outcomes and FVC was also associated with lower glucose and insulin.

Supplemental Figures 1-4 show that between SNP Wald ratio heterogeneity was high for most of the outcomes.

We found that genetic instruments for lung function were related to height, BMI, and education, with the strongest associations seen for height. There was no strong evidence of association with alcohol or tobacco phenotypes (Supplemental Table 4). Subsequent bi-Mendelian randomization analyses suggested there might be bi-directional effect between lung function and BMI, height, and education (Supplemental Table 5). Based on these findings we undertook multivariable MR adjusting for height only. This was done on the basis of the strong association of the lung function genetic instruments with height and knowledge that height is an established cause of lung function (Supplemental Table 4). As the associations of the genetic instruments for lung function with BMI and education were relatively weak and multivariable Mendelian randomization can introduce weak instrument bias we did not additionally adjust for these. Effect estimates from multivariable Mendelian randomization, adjusting for height, were directionally similar to the main IVW results. However, the results for CAD were attenuated quite notably ($OR_{MVMR-IVW}$ for 1SD FEV_1 0.95 (0.75 to 1.19) and for 1SD FVC 0.87 (0.70 to 1.09), though height adjustment did not attenuate effects for stroke or type 2 diabetes (Supplemental Figures 5-8). With adjustment for height the positive association of FEV_1 and FVC with atrial fibrillation disappeared whilst FEV_1 and FVC were potentially associated with lower risk of heart failure (Supplemental Figures 5 and 7). Similar findings were observed when we used multivariable MR-Egger method.

However, these analyses need to be treated with some caution as the conditional F statistics were small, ranging from 7.4 to 9.3 for FEV₁ and 9.7 to 11.4 for FVC.

Supplemental figures 1-4 show the MR-Egger and weighted median effect estimates of FEV₁ and FVC on CVD and cardiovascular risk factors. For FEV₁, the direction and magnitudes of the estimates from MR-Egger and weighted median were consistent with those from IVW, with the exception of large artery stroke. The MR-Egger analysis suggested that higher FEV₁ increased the risk of large artery stroke, albeit with wide confidence intervals overlapping null. However, the MR-Egger intercept test did not provide strong evidence for overall horizontal pleiotropy, except for atrial fibrillation. For FVC, the direction and magnitudes of the estimates from MR-Egger and weighted median were consistent with those from IVW. Similar findings to those seen from IVW analyses were observed when we used MR-PRESSO to control for pleiotropy by outlier exclusion (Supplemental Tables 6-7).

Figure 3 shows the association of predisposition to cardiovascular disease with FEV₁ and FVC, using genetic instruments for CVD (Supplemental Table 8). We found no strong evidence for an association between CVD predisposition and lung function.

Discussion

In this Mendelian randomization study, we have assessed the impact of lung function on the risk of CVDs and risk factors which may mediate any observed effects on CVDs (blood pressure, lipids, glycemic traits). We also undertook several sensitivity analyses to explore assumption violations and explored possible confounding by height and smoking. We found convincing evidence that lung function is

cardioprotective against CAD, consistent with our previous Mendelian randomization regarding the effect magnitude (OR_{FEV_1} : 0.78 and OR_{FVC} : 0.82).⁷ These findings were consistent across a range of sensitivity analyses used to explore possible bias due to horizontal pleiotropy, however they did attenuate notably with multivariable Mendelian randomization adjustment for height (to an OR of 0.95 for FEV_1 and 0.85 for FVC). We also found novel evidence that better lung function protects against stroke and type 2 diabetes, with these effects being consistent across sensitivity analyses, including adjustment for height. The positive association between lung function and atrial fibrillation in our main Mendelian randomization is in the opposite direction to that seen for other CVD outcomes but attenuated to the null with adjustment for height. The impact of lung function on heart failure is also less clear than that seen for CAD, stroke and type 2 diabetes, as the protective effect is only evident when adjusting for height in the multivariable Mendelian randomization. Lastly, our study also suggested that genetic susceptibility to CVD is unlikely to affect lung function.

Previous observational studies have suggested poorer FEV_1 and FVC increase CVD.^{1-5 35} Our study suggests these observed associations are likely causal for CAD and stroke. Multiple pathways may explain the protective effect, with our study suggesting systolic blood pressure is a possible mechanism,³⁶ but with weaker evidence of an effect of lung function on diastolic blood pressure. This may imply poorer lung function primarily acts on arterial stiffness.³⁷ We also found better lung function associated with lower triglycerides, lower insulin and type 2 diabetes. As previous Mendelian randomization studies support a causal effect of hypertriglyceridemia, hyperglycemia, type 2 diabetes and higher insulin on CAD,³⁸⁻⁴¹

these are possibly also mediators of the effect of lung function on CAD. With the exception of an inverse association with triglycerides, we did not find strong evidence for a causal effect of lung function on lipids. Importantly, we cannot rule out that the effect of lung function on CAD is largely explained by confounding due to height given the attenuation of the effect estimates in multivariable Mendelian randomization adjustment for height. These results have to be treated with some caution as the conditional F-statistics suggest possible weak instrument bias.

In our main analyses we found evidence of a protective effect of lung function on overall, ischemic, small vessel and large artery stroke, but not cardioembolic stroke. The underlying mechanisms may include reduced blood pressure,⁴² reduced risk of type 2 diabetes,⁴³ whilst the role of lipid is not as clear.⁴¹ The lack of large GWAS with publicly available data on haemorrhagic stroke mean that we were not able to explore this outcome and it is not included in the overall stroke outcome. With the exception of MR Egger results sensitivity analyses were largely supportive of these causal effects. The magnitude of the effect estimates for these stroke outcomes were smaller than those for CAD and confidence intervals were wider, suggesting that lung function may be a weaker risk factor for stroke than CAD. However, it is also possible that the weaker effect of lung function on stroke (compared with that observed for CAD) is an underestimation because of selection bias. For example, if lung function is also related to survival then those who survive to have a stroke (which tend to occur at older ages than CAD) will likely be the ones with better lung function. Stroke shares several other risk factors, such as hypertension with CAD, but commonly occurs after CAD and this competing risk may attenuate or even reverse the estimates.^{8 44} Similar paradoxical findings have also been observed in

observational studies, such as the association of cigarette smoking with dementia where the relation is attenuated, or even reversed with increasing age.⁴⁵ Further replication of our findings for stroke in Mendelian randomization studies ideally in cohorts with younger age at recruitment may better help elucidate the role of lung function in stroke. In contrast to CAD and stroke, results for heart failure in the main analysis were very close to the null, which could suggest lung function is not causally related to heart failure. However, it remains possible that the lack of association is due to biases arising from covariable adjustments on survival and/or competing risk of CAD.^{8 44}

The positive relation of lung function with atrial fibrillation in our main IVW analyses is somewhat unexpected given the majority of the findings from observational studies suggest an inverse relation.⁴⁶ However, the positive association disappeared upon adjustment for height using multivariable Mendelian randomization. Whilst the reason underpinning these changes is unclear, it is possible that height may be a stronger confounder for atrial fibrillation than any other forms of CVD, or adjustment for height may have reduced biases arising competing risk of CAD. This could also imply substantial horizontal pleiotropy due to height could have made other pleiotropy robust methods (e.g. MR-Egger, MR-PRESSO) less efficient which gave the same findings with IVW in the atrial fibrillation analyses. Nevertheless, the results from multivariable Mendelian randomization should be interpreted with caution due to wider confidence intervals and the possibility of weak instrument bias.

Although we used Mendelian randomization which is less susceptible to confounding by key traits, such as smoking, height and socio-economic position,⁴⁷ there are some

limitations. First, the validity of our study depends on the 3 main instrumental variable assumptions.⁶ The F-statistics and proportion of variation in FEV₁ and FVC make a major impact of weak instrument bias on our main IVW results unlikely. We chose SNPs identified from the UK Biobank summary statistics as instruments for lung function. Whilst this avoids the introduction of invalid SNPs due to bias arising from adjustment for height and smoking (as done in previous GWAS),⁴⁸ the lack of replication of the UK Biobank GWAS results may mean that some of the genetic instrument-FEV₁ and FVC associations are exaggerated and hence the Mendelian randomization results presented here attenuated. However, the broad similarity of our findings here for CAD with those from our previous paper (using GWAS summary data from a collaboration of studies not only including UK Biobank) suggest this has not caused major bias (OR_{FEV1}: 0.78 and OR: 0.82_{FVC} in our previous study versus OR: 0.72_{FEV1} and OR: 0.70_{FVC} in this study).⁷ We also explored whether the lung function genetic instruments were associated with potential confounders of the relation of lung function with CVDs risk (i.e. BMI, smoking, alcohol and education, and height), which if present could bias our results through horizontal pleiotropy. We found limited evidence that the lung function instruments relate strongly to these risk factors, apart from height. Multivariable Mendelian randomization to adjust for height did notably attenuate the effects on CAD, though not those for stroke or type 2 diabetes. However, additional exploratory analyses suggested the magnitude, but not direction, of some multivariable Mendelian randomization estimates were sensitive to the choice of GWAS where height instruments were extracted from, such as directly from GIANT (Supplemental Figures 5-8) or UK Biobank (Supplemental Figures 9-12). As such, interpretation of these adjusted results is difficult because of their imprecision and the possibility of

weak instrument bias and it would be valuable to repeat our analyses using data from even larger summary GWAS. Whilst we found evidence for between SNP heterogeneity, which might occur because of horizontal pleiotropy resulting in violation of the third assumption, sensitivity analyses that explore bias due to unbalanced horizontal pleiotropy (i.e. MR-Egger, weighted median and MR-PRESSO analyses) supported our main analyses. In addition to these three key assumptions, two-sample MR assumes that the two samples are from the same underlying population and that the two samples are independent. With the exception of summary data for CAD (23% non-European), all of our samples were restricted to European populations. There was overlap between the lung function and blood pressure (100% - i.e. both were undertaken in UK biobank), atrial fibrillation (38% of this outcome GWAS participants were in UK biobank), and heart failure (40% in UK biobank). Such overlap can mean that any weak instrument bias could be towards confounded estimates, rather than towards the null (as in two sample MR with independent samples). This would particularly influence blood pressure and could mean associations are exaggerated by in the presence of weak instruments. Second, we were unable to explore the possibility of non-linear effects between lung function and CVD, which can only be explored in one sample Mendelian randomization in large biobanks with individual level data.⁴⁹

In conclusion, our Mendelian randomization study provides some evidence concerning the protective role of lung function on CAD, stroke, type 2 diabetes and lower systolic blood pressure. Future studies should explore the underlying mechanisms, including the role of height in these relationships, and hence help identify additional targets of intervention for cardiovascular disease prevention.

Future GWAS and Mendelian randomization studies exploring causes of hemorrhagic stroke are also important, particularly for settings where hemorrhagic stroke is prevalent, such as China.⁵⁰

Acknowledgement

Data on coronary artery disease have been contributed by CARDIoGRAMplusC4D investigators and have been downloaded from www.CARDIORAMPLUSC4D.ORG. We thank the investigators of the atrial fibrillation GWAS for providing the summary statistics, which have been downloaded from <http://csg.sph.umich.edu/willer/public/afib2018/>. We thank MEGASTROKE for providing the summary statistics for stroke, which have been downloaded from <http://www.megastroke.org/index.html>. The MEGASTROKE project received funding from sources specified at <http://www.megastroke.org/acknowledgments.html>. The list of authors in the MEGASTROKE Consortium can be found in the Supplemental materials. We thank HERMES Consortium for providing the summary statistics for heart failure, which have been downloaded from <http://www.broadcvdi.org/informational/data>. We thank GLGC for providing summary statistics for lipids, which have been downloaded from <http://csg.sph.umich.edu/willer/public/lipids2013/>. We thank MAGIC for providing summary statistics for glycemic traits, which have been downloaded from <https://www.magicinvestigators.org>. We thank DIAGRAM for providing the summary statistics for type 2 diabetes, which have been downloaded from <https://www.diagram-consortium.org>. We thank GIANT consortium for providing the summary statistics for body mass index and height, which have been downloaded from

https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files. We thank SSGAC for providing the summary statistics for years of education attainment, which have been downloaded from <https://www.thessgac.org/data>. We thank GSCAN for providing the summary statistics for alcohol and cigarette use phenotypes, which have been downloaded from <https://conservancy.umn.edu/handle/11299/201564>. Data on FEV₁ and FVC, systolic and diastolic blood pressure were extracted from summary genome-wide association study results in the UK Biobank, available in the IEU GWAS database (<https://gwas.mrcieu.ac.uk/>). We thank Mr Wong KN for summarizing the descriptive information regarding the GWAS used in this study.

Disclosure

DAL receives support from several national and international government and charitable research funders, as well as from Medtronic Ltd and Roche Diagnostics for research unrelated to that presented here. All other authors declare they have no conflict of interest, financial or otherwise.

Authors' contribution

SLAY designed the study, wrote the analysis plan and interpreted the results. SLAY undertook analyses with feedback from MCB, DAL, and CMS. SLAY wrote the first draft of the manuscript with critical feedback and revisions from MCB, DAL, and CMS. All authors gave final approval of the version to be published. SLAY had primary responsibility for final content.

Funding

MCB is supported by MRC Skills Development Fellowship (MR/P014054/1). MCB and DAL's contribution to this study is supported by the British Heart Foundation (AA/18/7/34219) and MCB and DAL work in a Unit receives funding from the University of Bristol and UK Medical Research Council (MRC) (MC_UU_00011/6). The funders had no role in the design, analyses, interpretation of results or writing of the paper. The views expressed in this paper are those of the authors and not necessarily any of the funders.

References

1. Silvestre OM, Nadruz W, Jr., Querejeta Roca G, et al. Declining Lung Function and Cardiovascular Risk: The ARIC Study. *J Am Coll Cardiol* 2018;72(10):1109-22. doi: 10.1016/j.jacc.2018.06.049 [published Online First: 2018/09/01]
2. Engstrom G, Melander O, Hedblad B. Population-based study of lung function and incidence of heart failure hospitalisations. *Thorax* 2010;65(7):633-8. doi: 10.1136/thx.2010.135392 [published Online First: 2010/07/16]
3. Hozawa A, Billings JL, Shahar E, et al. Lung function and ischemic stroke incidence: the Atherosclerosis Risk in Communities study. *Chest* 2006;130(6):1642-9. doi: 10.1378/chest.130.6.1642 [published Online First: 2006/12/15]
4. Johnston AK, Mannino DM, Hagan GW, et al. Relationship between lung function impairment and incidence or recurrence of cardiovascular events in a middle-aged cohort. *Thorax* 2008;63(7):599-605. doi: 10.1136/thx.2007.088112 [published Online First: 2008/02/05]
5. Sin DD, Wu LL, Man SFP. The relationship between reduced lung function and cardiovascular mortality - A population-based study and a systematic review of the literature. *Chest* 2005;127(6):1952-59.
6. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;362:k601. doi: 10.1136/bmj.k601 [published Online First: 2018/07/14]
7. 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(4):e001952. doi: 10.1161/CIRCGEN.117.001952 [published Online First: 2018/04/14]
8. Hartwig FP, Tilling K, Davey Smith G, et al. Bias in two-sample Mendelian randomization by using covariable-adjusted summary associations. *bioRxiv* 2019 doi: <https://doi.org/10.1101/816363>
9. 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(1):119. doi: 10.1038/s42003-019-0361-2 [published Online First: 2020/01/12]

10. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol* 2016;45(3):908-15. doi: 10.1093/ije/dyw127
11. Yang Q, Sanderson E, Tilling K, et al. Exploring and mitigating potential bias when genetic instrumental variables are associated with multiple non-exposure traits in Mendelian randomization. *medRxiv* 2019:19009605. doi: 10.1101/19009605
12. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;47(10):1121-30. doi: 10.1038/ng.3396
13. 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(4):524-37. doi: 10.1038/s41588-018-0058-3 [published Online First: 2018/03/14]
14. Nielsen JB, Thorolfsson RB, Fritsche LG, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat Genet* 2018;50(9):1234-39. doi: 10.1038/s41588-018-0171-3 [published Online First: 2018/08/01]
15. Shah S, Henry A, Roselli C, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. *Nat Commun* 2020;11(1):163. doi: 10.1038/s41467-019-13690-5 [published Online First: 2020/01/11]
16. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12(3):e1001779. doi: 10.1371/journal.pmed.1001779 [published Online First: 2015/04/01]
17. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018;7 doi: 10.7554/eLife.34408 [published Online First: 2018/05/31]
18. Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013;45(11):1274-83. doi: 10.1038/ng.2797
19. Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet* 2019;51(2):237-44. doi: 10.1038/s41588-018-0307-5 [published Online First: 2019/01/16]
20. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;518(7538):197-206. doi: 10.1038/nature14177
21. Manning AK, Hivert MF, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet* 2012;44(6):659-69. doi: 10.1038/ng.2274 [published Online First: 2012/05/15]
22. Wheeler E, Leong A, Liu CT, et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS Med* 2017;14(9):e1002383. doi: 10.1371/journal.pmed.1002383 [published Online First: 2017/09/13]
23. Scott RA, Scott LJ, Magi R, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* 2017 doi: 10.2337/db16-1253

24. Okbay A, Beauchamp JP, Fontana MA, et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016;533(7604):539-42. doi: 10.1038/nature17671
25. Wood AR, Esko T, Yang J, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 2014;46(11):1173-86. doi: 10.1038/ng.3097
26. Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *International Journal of Epidemiology* 2011;40(3):740-52. doi: 10.1093/ije/dyq151 [published Online First: 2010/09/04]
27. Yarmolinsky J, Bonilla C, Haycock PC, et al. Circulating Selenium and Prostate Cancer Risk: A Mendelian Randomization Analysis. *J Natl Cancer Inst* 2018;110(9):1035-38. doi: 10.1093/jnci/djy081 [published Online First: 2018/05/23]
28. Bowden J, Del Greco MF, Minelli C, et al. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med* 2017;36(11):1783-802. doi: 10.1002/sim.7221 [published Online First: 2017/01/24]
29. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* 2015;181(4):251-60. doi: 10.1093/aje/kwu283
30. 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(3):713-27. doi: 10.1093/ije/dyy262 [published Online First: 2018/12/12]
31. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med* 2017;36(29):4705-18. doi: 10.1002/sim.7492 [published Online First: 2017/09/30]
32. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44(2):512-25. doi: 10.1093/ije/dyv080
33. Bowden J, Davey Smith G, Haycock PC, et al. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* 2016;40(4):304-14. doi: 10.1002/gepi.21965
34. Verbanck M, Chen CY, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 2018;50(5):693-98. doi: 10.1038/s41588-018-0099-7 [published Online First: 2018/04/25]
35. Duong M, Islam S, Rangarajan S, et al. Mortality and cardiovascular and respiratory morbidity in individuals with impaired FEV1 (PURE): an international, community-based cohort study. *Lancet Glob Health* 2019;7(5):e613-e23. doi: 10.1016/S2214-109X(19)30070-1 [published Online First: 2019/04/20]
36. Jacobs DR, Jr., Yatsuya H, Hearst MO, et al. Rate of decline of forced vital capacity predicts future arterial hypertension: the Coronary Artery Risk Development in Young Adults Study. *Hypertension* 2012;59(2):219-25. doi: 10.1161/HYPERTENSIONAHA.111.184101 [published Online First: 2011/12/29]
37. O'Rourke M. Arterial stiffness, systolic blood pressure, and logical treatment of arterial hypertension. *Hypertension* 1990;15(4):339-47. [published Online First: 1990/04/01]

38. Zhao JV, Luo S, Schooling CM. Sex-specific Mendelian randomization study of genetically predicted insulin and cardiovascular events in the UK Biobank. *Commun Biol* 2019;2:332. doi: 10.1038/s42003-019-0579-z [published Online First: 2019/09/12]
39. Ahmad OS, Morris JA, Mujammami M, et al. A Mendelian randomization study of the effect of type-2 diabetes on coronary heart disease. *Nat Commun* 2015;6:7060. doi: 10.1038/ncomms8060 [published Online First: 2015/05/29]
40. Au Yeung SL, Luo S, Schooling CM. The Impact of Glycated Hemoglobin (HbA1c) on Cardiovascular Disease Risk: A Mendelian Randomization Study Using UK Biobank. *Diabetes Care* 2018;41(9):1991-97. doi: 10.2337/dc18-0289 [published Online First: 2018/06/29]
41. Allara E, Morani G, Carter P, et al. Genetic Determinants of Lipids and Cardiovascular Disease Outcomes: A Wide-Angled Mendelian Randomization Investigation. *Circ Genom Precis Med* 2019;12(12):e002711. doi: 10.1161/CIRCGEN.119.002711 [published Online First: 2019/11/23]
42. Flint AC, Conell C, Ren X, et al. Effect of Systolic and Diastolic Blood Pressure on Cardiovascular Outcomes. *N Engl J Med* 2019;381(3):243-51. doi: 10.1056/NEJMoa1803180 [published Online First: 2019/07/18]
43. Larsson SC, Scott RA, Traylor M, et al. Type 2 diabetes, glucose, insulin, BMI, and ischemic stroke subtypes: Mendelian randomization study. *Neurology* 2017;89(5):454-60. doi: 10.1212/WNL.0000000000004173 [published Online First: 2017/07/02]
44. Schooling C, Lopez P, Yang Z, et al. Bias from competing risk before recruitment in Mendelian Randomization studies of conditions with shared etiology. *bioRxiv* 2020
45. Hernan MA, Alonso A, Logroscino G. Cigarette smoking and dementia: potential selection bias in the elderly. *Epidemiology* 2008;19(3):448-50. doi: 10.1097/EDE.0b013e31816bbe14 [published Online First: 2008/04/17]
46. Li J, Agarwal SK, Alonso A, et al. Airflow obstruction, lung function, and incidence of atrial fibrillation: the Atherosclerosis Risk in Communities (ARIC) study. *Circulation* 2014;129(9):971-80. doi: 10.1161/CIRCULATIONAHA.113.004050
47. Davey Smith G, Lawlor DA, Harbord R, et al. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med* 2007;4(12):e352. doi: 10.1371/journal.pmed.0040352 [published Online First: 2007/12/14]
48. Soler Artigas M, Wain LV, Miller S, et al. Sixteen new lung function signals identified through 1000 Genomes Project reference panel imputation. *Nat Commun* 2015;6:8658. doi: 10.1038/ncomms9658
49. Sun YQ, Burgess S, Staley JR, et al. Body mass index and all cause mortality in HUNT and UK Biobank studies: linear and non-linear mendelian randomisation analyses. *BMJ* 2019;364:l1042. doi: 10.1136/bmj.l1042 [published Online First: 2019/04/09]
50. Feigin VL, Roth GA, Naghavi M, et al. Global burden of stroke and risk factors in 188 countries, during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet Neurol* 2016;15(9):913-24. doi: 10.1016/S1474-4422(16)30073-4

Table 1: Data sources used in this Mendelian randomization study

Exposure	Data source (PMID)	Sample size (% cases)	% European
Forced expiratory volume in 1 second (SD)	UK Biobank (IEU GWAS)	421,986	100
Forced vital capacity (SD)	UK Biobank (IEU GWAS)	421,986	100
Outcomes	Data source (PMID)	Sample size (% cases)	% European
Coronary artery disease	CARDIoGRAM (26343387)	184,305 (33%)	77
Stroke	MEGASTROKE (29531354)	446,696 (9%)	100
Ischemic stroke	MEGASTROKE (29531354)	440,328 (8%)	100
Cardioembolic stroke	MEGASTROKE (29531354)	211,763 (3%)	100
Small vessel stroke	MEGASTROKE (29531354)	198,048 (3%)	100
Large artery stroke	MEGASTROKE (29531354)	150,765 (3%)	100
Atrial fibrillation	PMID: 30061737	1,030,836 (6%)	100
Heart failure	HERMES (31919418)	977,320 (5%)	100
Systolic blood pressure (SD)	UK Biobank (IEU GWAS)	436,419	100
Diastolic blood pressure (SD)	UK Biobank (IEU GWAS)	436,424	100
LDL cholesterol (SD)	GLGC (24097068)	173,082	100
HDL cholesterol (SD)	GLGC (24097068)	187,167	100
Triglycerides (SD)	GLGC (24097068)	177,861	100
Glucose (mmol/L)	MAGIC (22581228)	58,074	100
HbA1c (%)	MAGIC (28898252)	123,665	100
Insulin (log)	MAGIC (22581228)	51,750	100
Type 2 diabetes	DIAGRAM (28566273)	159,208 (17%)	100
Confounders	Data source (PMID)	Sample size	% European
Body mass index (SD)	GIANT (25673413)	339,224	95
Years education attained (SD)	SSGAC (27225129)	328,917	100
Alcohol (SD of Log transformed drinks per week)	GSCAN (30642351)	941,280	100
Smoking Heaviness (SD of cigarettes per day)	GSCAN (30642351)	337,334	100
Height (SD)	GIANT (25282103)	253,288	100

Figure 1: The impact of forced expiratory volume in 1 second (FEV₁, per SD) and forced vital capacity (FVC, per SD) on cardiovascular disease and type 2 diabetes using Mendelian randomization

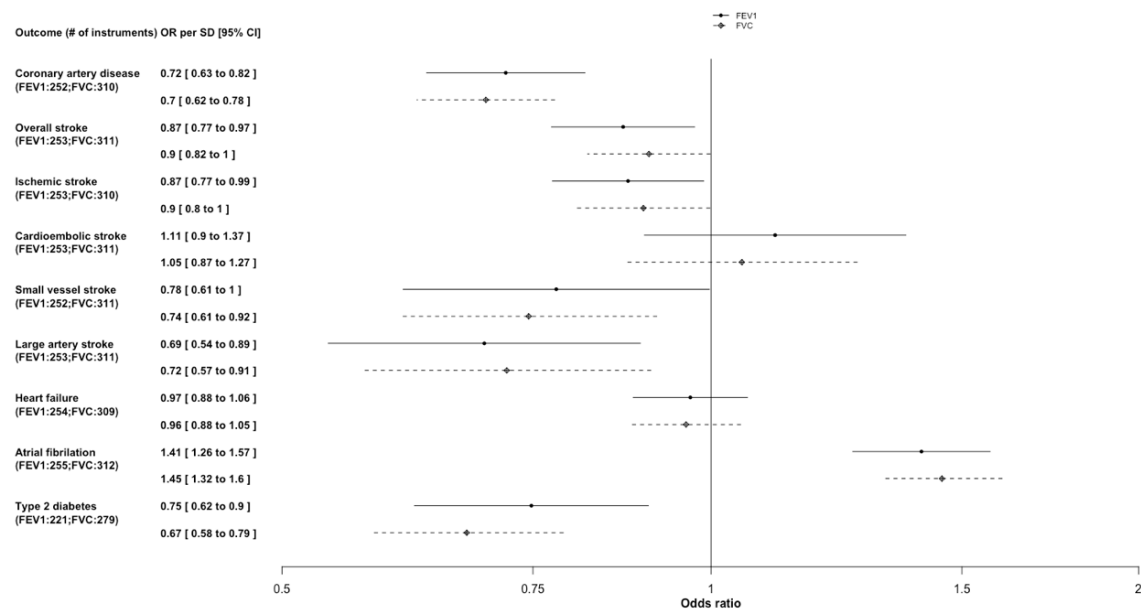


Figure 2: The impact of forced expiratory volume in 1 second (FEV₁, per SD) and forced vital capacity (FVC, per SD) on continuously measured cardiovascular risk factors using Mendelian randomization

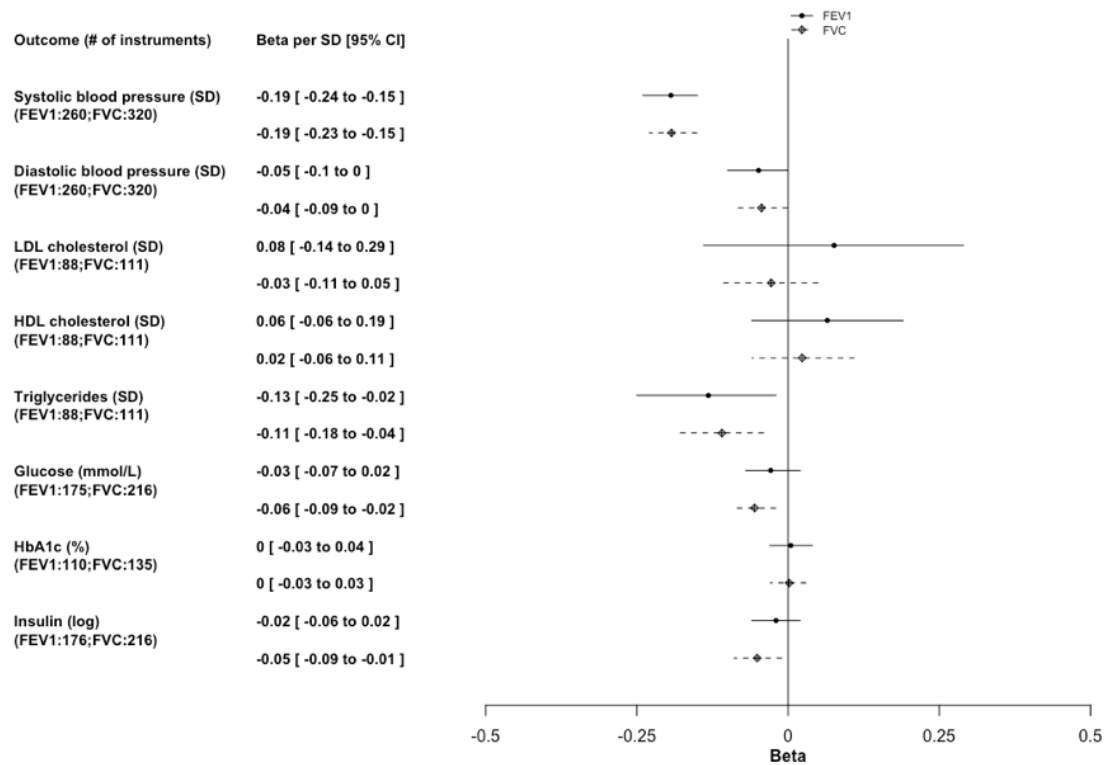


Figure 3: The impact of higher predisposition to cardiovascular outcome on forced expiratory volume in 1 second (FEV₁) and forced vital capacity (FVC) using Mendelian randomization

