Au Yeung, S. L. R., Borges, M. C., Lawlor, D. A., & Schooling, C. M. (2021). Impact of lung function on cardiovascular diseases and cardiovascular risk factors: A two sample bidirectional Mendelian randomisation study. *Thorax*, [215600]. https://doi.org/10.1136/thoraxjnl-2020-215600

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

Shiu Lun Au Yeung[1*], Maria Carolina Borges[2,3], Debbie A Lawlor[2,3,4], C Mary Schooling[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: 250

Main text: 3,414

Tables: 1

Figures: 5

Supplemental material: 1

Running title: Lung function and CVD using MR

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

Abstract

Introduction: Observational studies suggested lung function is inversely associated with cardiovascular disease (CVD) although these studies could be confounded. We conducted a two 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 (FEV_1: 260) and forced vital capacity (FVC: 320) from publicly available UK Biobank summary statistics (n=421,986) and applied to GWAS summary statistics for coronary artery disease (CAD) (n=184,305), stroke (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, with various sensitivity analyses. Bi-directional Mendelian randomization was used to assess reverse causation.

Results: FEV_1 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)), and some stroke subtypes. FEV_1 and FVC were inversely associated with type 2 diabetes and systolic blood pressure. Sensitivity analyses produced similar findings although the association with CAD was attenuated after adjusting for height (e.g. OR for 1SD FEV_1 0.95 (0.75 to 1.19), but not for stroke or type 2 diabetes. There was no strong evidence for reverse causation.

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

Key messages

What is the key question?

Does higher lung function protect against cardiovascular disease, and is there evidence of reverse causation?

What is the bottom line?

This Mendelian randomization study suggests higher lung function likely reduces coronary artery disease and stroke risk, and there is little evidence suggesting reverse causation.

Why read on?

This is one of the largest Mendelian randomization studies exploring the impact of lung function in a wide range of cardiovascular diseases and their risk factors, as well as assessing possible reverse causation.

Introduction

Poorer lung function is associated with cardiovascular disease (CVD).[1-4] 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.[5] A previous Mendelian randomization study suggested higher FEV_1 related to lower risk of coronary artery disease (CAD), whilst the relation for FVC was less clear.[6] 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).[6] Furthermore, the study used genetic instruments extracted from genome wide association studies which were adjusted for height and smoking, which may introduce collider bias and hence potentially identify invalid instruments.[7] Another Mendelian randomization study provided evidence that lung function may protect against CAD although that study focused primarily on the mechanistic pathway between height and CAD risk instead of the etiologic role of lung function in cardiovascular diseases.[8]

The aim of this study was to determine the causal effect of FEV_1 and FVC on a wide range of cardiovascular diseases using two sample Mendelian randomization with summary statistics.[5] 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 FEV_1 (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).[5] Table 1 summarizes the study design including the sources of GWAS summary data for exposure (FEV_1 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.[9]

Ethics approval

In this study we have used publicly-available GWAS results from relevant publications and database (https://gwas.mrcieu.ac.uk/).[10-23] 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 FEV_1 and FVC (Sample 1)

Genetic determinants of FEV_1 (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/).[14 15] 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),[10] stroke and its subtypes (MEGASTROKE consortium),[11] atrial fibrillation,[12] heart failure (HERMES consortium),[13] systolic and diastolic blood pressure (UK Biobank associations obtained from IEU GWAS database),[14 15] low density lipoprotein (LDL), high density lipoprotein (HDL) cholesterol and triglycerides (GLGC),[16] type 2 diabetes (DIAGRAM),[21] fasting glucose, insulin, and glycated

hemoglobin (MAGIC),[19 20] were obtained from publicly available online GWAS summary data repositories, with majority of them retrieved via MR Base.[15] 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 ($<5x10^{-8}$), 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).[15] 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 (values outside of 0.42 to 0.58) 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.[24] For our main Mendelian randomization analyses, we used inverse variance weighting (IVW) with multiplicative random effects to obtain the causal effect of FEV_1 and FVC on CVD outcomes and their risk factors. IVW assumes there is no unbalanced horizontal pleiotropy.[25] 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 Cochrane 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 FEV_1 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).[9] This was done using publicly available summary genome wide association study data for: body mass index,[18] education (as a measure of socioeconomic position),[22] number of drinks of alcohol per week and number of cigarette smoked per day,[17] and height.[23] 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 were 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.[9] 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.[9] Given the possibility of varying instrument strengths with increasing number of variables being adjusted for in multivariable Mendelian randomization and hence limits the interpretation of the findings,[26] we only considered variables which were strongly associated with the instruments for FEV_1 and FVC. We also approximated the conditional F statistics to evaluate potential weak instrument bias,[26] presented Cochrane Q test to evaluate heterogeneity, and also provided estimates for multivariable MR-Egger to assess robustness of findings due to pleiotropy.[27]

To assess unbalanced horizontal pleiotropy bias we undertook the following sensitivity analyses: MR-Egger,[28] weighted median analyses,[29] and MR-PRESSO,[30] 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). 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 FEV_1 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 FEV_1 or FVC , we also conducted a Mendelian randomization using genetic predictors of risk of CAD, overall stroke, heart failure, and atrial fibrillation as instruments.[10-13] We used IVW to explore evidence of effects of genetic predispositions to these CVDs on FEV_1 or FVC.

All analyses were performed using R Version 3.5.2 (R Development Core Team, Vienna, Austria) using R packages ("TwoSampleMR"), and ("MRPRESSO).[15 30]

Results

For FEV_1, up to 260 SNPs were used in the analyses, and these instruments explained up to 3.5% of FEV_1 variance (Overall F statistics: 58.8). For FVC, up to 320 SNPs were used in the analyses, and these instruments explained up to 4.8% of FVC variance (Overall F statistics: 66.4). Supplemental Tables 2-3 shows the summary statistics of the lung function instruments used in this study.

Figure 1 shows that higher FEV_1 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 and had no association with heart failure. FVC showed similar relation with these outcomes (Figure 2). Regarding the association with cardiovascular risk factors, higher FEV_1 was mainly associated with lower systolic blood pressure and triglycerides (Figure 3). FVC showed similar relations with these outcomes and FVC was also associated with lower LDL cholesterol, glucose and insulin (Figure 4). Figures 1-4 show that

between SNP Wald ratio heterogeneity was high for most of the outcomes although there was no strong evidence for directional horizontal pleiotropy based on the MR-Egger intercepts, with the exception of atrial fibrillation. Corresponding sensitivity analyses generally gave directionally similar estimates for most outcomes although differences were observed when using multivariable Mendelian randomization analyses (Figures 1-4, and Supplemental Tables 4-5).

Based on the results in Supplemental Tables 6-7, we identified BMI, height and education as possible horizontal pleiotropic effects which may bias the main analyses. Given height had a much stronger association with the instruments and knowledge that height is an established cause of lung function, we undertook multivariable MR adjusting for height only. With adjustment for height, the effect estimates were directionally similar to the main analyses although estimates for CAD attenuated substantially, whilst the effect estimate for stroke and type 2 diabetes remained similar. Given the lower statistical power in these analyses the confidence intervals are wider, as expected, and in the case of stroke included the null value (Figures 1 and 2). The positive association with atrial fibrillation disappeared and lung function was potentially associated with lower risk of heart failure although with confidence interval overlapping null (Figures 1 and 2). Similar findings were observed when we used multivariable MR-Egger method (Supplemental Figures 1-4). However, these analyses need to be treated with some caution as the conditional F statistics were small, ranging from 7.8 to 9.9 for FEV_1 and 10.0 to 11.6 for FVC.

Figure 5 shows the association of predisposition to cardiovascular disease with FEV_1 and FVC, using genetic instruments for CVD (Supplemental Table 8). Predisposition

to coronary artery disease, stroke, atrial fibrillation and heart failure was not associated with FEV_1 or FVC. This provides evidence against possible reverse causation when assessing the relation of lung function with cardiovascular diseases.

Discussion

In this Mendelian randomization study which we explored 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). Our study provides suggestive evidence that better lung function is cardioprotective. These findings were consistent across a range of sensitivity analyses used to explore possible bias due to horizontal pleiotropy, except for CAD where the association was attenuated after adjusting for height. 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.[1-4] 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, 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,[31] and hence may contribute to systemic hypertension. We also found better lung function associated with lower triglycerides and LDL cholesterol, lower insulin and type 2 diabetes. As previous Mendelian randomization studies support a causal effect of hypercholesterolemia, hyperglycemia, type 2 diabetes and higher insulin on CAD,[32-35] 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. Our study also indicated the importance of height as a horizontal pleiotropic effect, given the associations with coronary artery disease were attenuated in the multivariable Mendelian randomization. Nevertheless, further investigation would be needed as attenuation may reflect weak instrument bias and was sensitive to where the SNP-height estimates in the multivariable Mendelian randomization were extracted from (GIANT in Supplemental Figure 1 versus UK Biobank in Supplemental Figure 5).

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,[36] reduced risk of type 2 diabetes,[37] whilst the role of lipid is not as clear.[35] The lack of large GWAS with publicly available data on hemorrhagic stroke mean that we were not able to explore this outcome and it is not included in the overall stroke outcome. Sensitivity analyses were largely supportive of these causal effects, apart for MR-Egger for some stroke subtypes. Although the estimates for stroke were smaller than those for

CAD and confidence intervals were wider, the apparently weaker effect of lung function on stroke (compared with that observed for CAD) could be 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.[7 38] 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.

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 although the association disappeared after adjusting for height.[39] Whilst the reason underpinning these changes is unclear, a recent Mendelian randomization study suggested taller people had higher risk of atrial fibrillation and hence failure to take into account height in the analyses may have led to the positive observed association in the main analyses given some of the lung function instruments were also related to height.[40] This may also explain the null findings upon adjustment for height in the multivariable Mendelian randomization.

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, after adjusting for height, FEV_1 and FVC were possibly associated with lower risk of heart failure (Figures 1 and 2). Although this could be a reflection of

adjustment removing possible selection bias, reasons underpinning the differences would require further investigation**.**

Our study, using a design that is less likely to be confounded than conventional multivariable regression, indicates a possible role of lung function on cardiovascular health. From a clinical perspective, our results support further exploration of the effectiveness of monitoring cardiovascular and type 2 diabetes risk in patients with poor lung function. They also support interventions which improve lung function, such as tobacco cessation and increased physical activity, to prevent cardiovascular disease.

We have not been able to fully explore the relation of the FEV_1/FVC ratio (a useful diagnostic marker of chronic obstructive pulmonary diseases)[41] with cardiovascular diseases or its risk factors. Nevertheless, exploratory analyses, using up to 97 SNPs strongly (p value $<5x10^{-8}$ and main reported trait) and independently ($r^2<0.001$) associated with FEV_1/FVC extracted from a recent GWAS (F statistic: 144),[42] suggested no strong evidence for an effect of FEV_1/FVC on any of the outcomes (Supplemental Figures 9-10). However, the interpretation of the findings for FEV_1/FVC ratio is potentially more challenging given any change in the ratio may also have reflected the varying magnitude of changes of both FEV_1 and FVC, which could have explained the differences with our main findings.

Although we used Mendelian randomization which is less susceptible to confounding by key traits, such as smoking, height and socio-economic position, there are some limitations. One of the limitations is that the validity of our study depends on the 3

main instrumental variable assumptions.[5] The F-statistics and proportion of variation in FEV_1 and FVC make a major impact of weak instrument bias on our main IVW results unlikely, which if present, would bias our estimates towards the null for majority of the outcomes. Previous studies suggested possible biases related to the use of Wald ratio to estimate causal effects for disease outcomes.[43] However, we expected such biases would be small given the large sample sizes used. We also explored associations of instruments with confounders (e.g. smoking and height) which can constitute violation of the assumptions, where we found height as a main pleiotropic effect. Although consistent findings were observed using standard sensitivity analyses, we noticed attenuations of estimates for coronary artery disease upon adjusting for height in the multivariable Mendelian randomization. However, the degree of attenuation depends on the choice of GWAS where height instruments were extracted from (Supplemental Figures 1-8), which could possibly be a reflection of weak instrument bias and warrant further investigations. Another limitation is that 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. A final limitation is that we were unable to disentangle the independent effects of FEV_1 and FVC on the outcomes. Whist exploratory analyses using multivariable Mendelian randomization suggested FVC may be more relevant than FEV_1 (Supplemental Figures 11-12), these results are susceptible to weak instrument bias (conditional F statistics <2.1 for both exposures) and so should be interpreted with caution.[26]

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.

Acknowledgements

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 FEV_1 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 KN Wong and Ms W Liang for summarizing the descriptive information regarding the GWAS used in this study. We thank Dr Eleanor Sanderson for helpful discussion concerning the multivariable Mendelian randomization.

Sources of 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.

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.

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. 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.
5. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol* 2016;45(3):908-15. doi: 10.1093/ije/dyw127
6. 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]
7. 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

8. 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]
9. 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
10. 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
11. 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]
12. Nielsen JB, Thorolfsdottir 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]
13. 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]
14. 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]
15. 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]
16. 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
17. 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]
18. 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
19. 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]
20. 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]

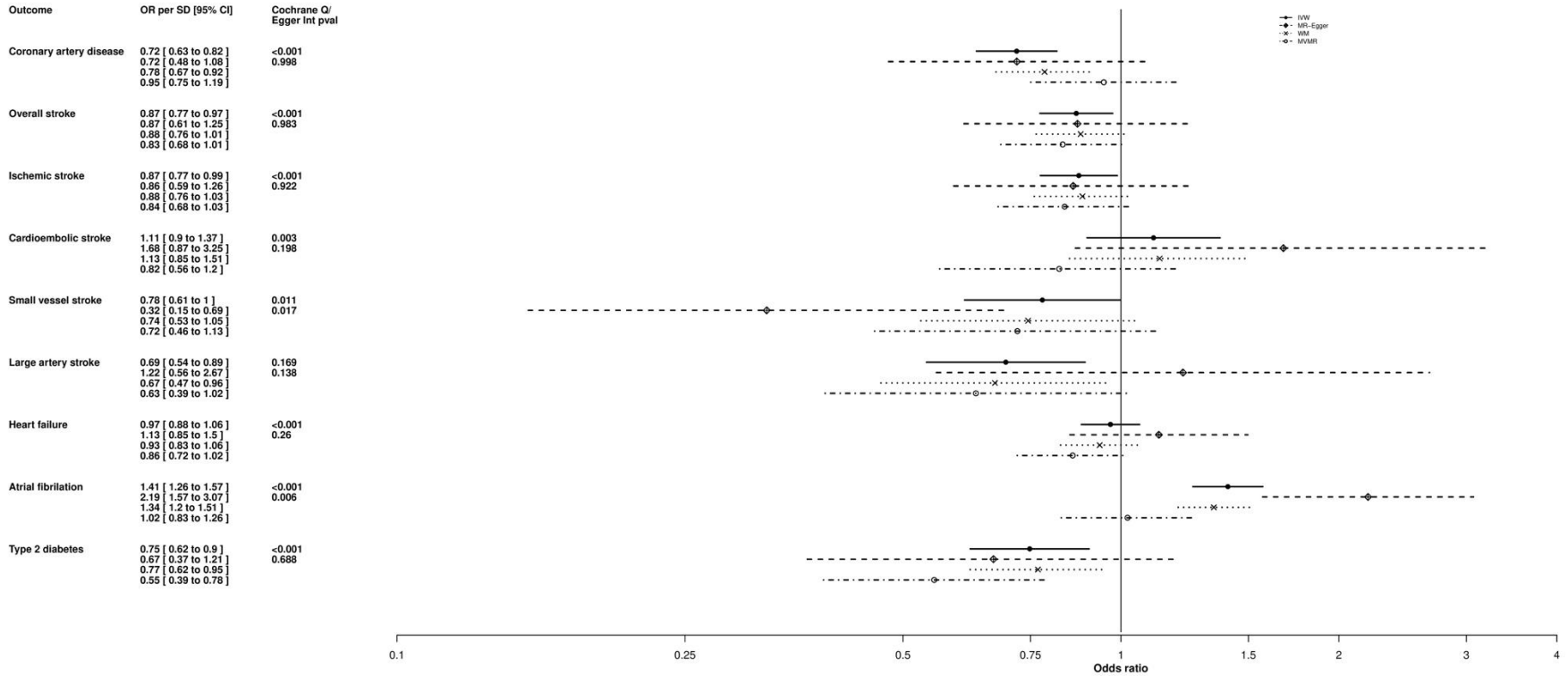
21. Scott RA, Scott LJ, Magi R, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* 2017 doi: 10.2337/db16-1253
22. Okbay A, Beauchamp JP, Fontana MA, et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016;533(7604):539-42. doi: 10.1038/nature17671
23. 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
24. 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]
25. 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]
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;48(3):713-27. doi: 10.1093/ije/dyy262 [published Online First: 2018/12/12]
27. 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]
28. 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
29. 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
30. 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]
31. 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]
32. 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]
33. 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]
34. 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]
35. 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]
36. 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]
37. 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]
38. Schooling C, Lopez P, Yang Z, et al. Use of multivariable Mendelian randomization to address biases due to competing risk before recruitment *Frontiers in Genetics* 2021
39. 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
40. Levin MG, Judy R, Gill D, et al. Genetics of height and risk of atrial fibrillation: A Mendelian randomization study. *PLoS Med* 2020;17(10):e1003288. doi: 10.1371/journal.pmed.1003288 [published Online First: 2020/10/09]
41. Johns DP, Walters JA, Walters EH. Diagnosis and early detection of COPD using spirometry. *J Thorac Dis* 2014;6(11):1557-69. doi: 10.3978/j.issn.2072-1439.2014.08.18 [published Online First: 2014/12/06]
42. 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(3):481-93. doi: 10.1038/s41588-018-0321-7 [published Online First: 2019/02/26]
43. Harbord RM, Didelez V, Palmer TM, et al. Severity of bias of a simple estimator of the causal odds ratio in Mendelian randomization studies. *Stat Med* 2013;32(7):1246-58. doi: 10.1002/sim.5659 [published Online First: 2012/10/20]

Table 1: Data sources used in this Mendelian randomization study

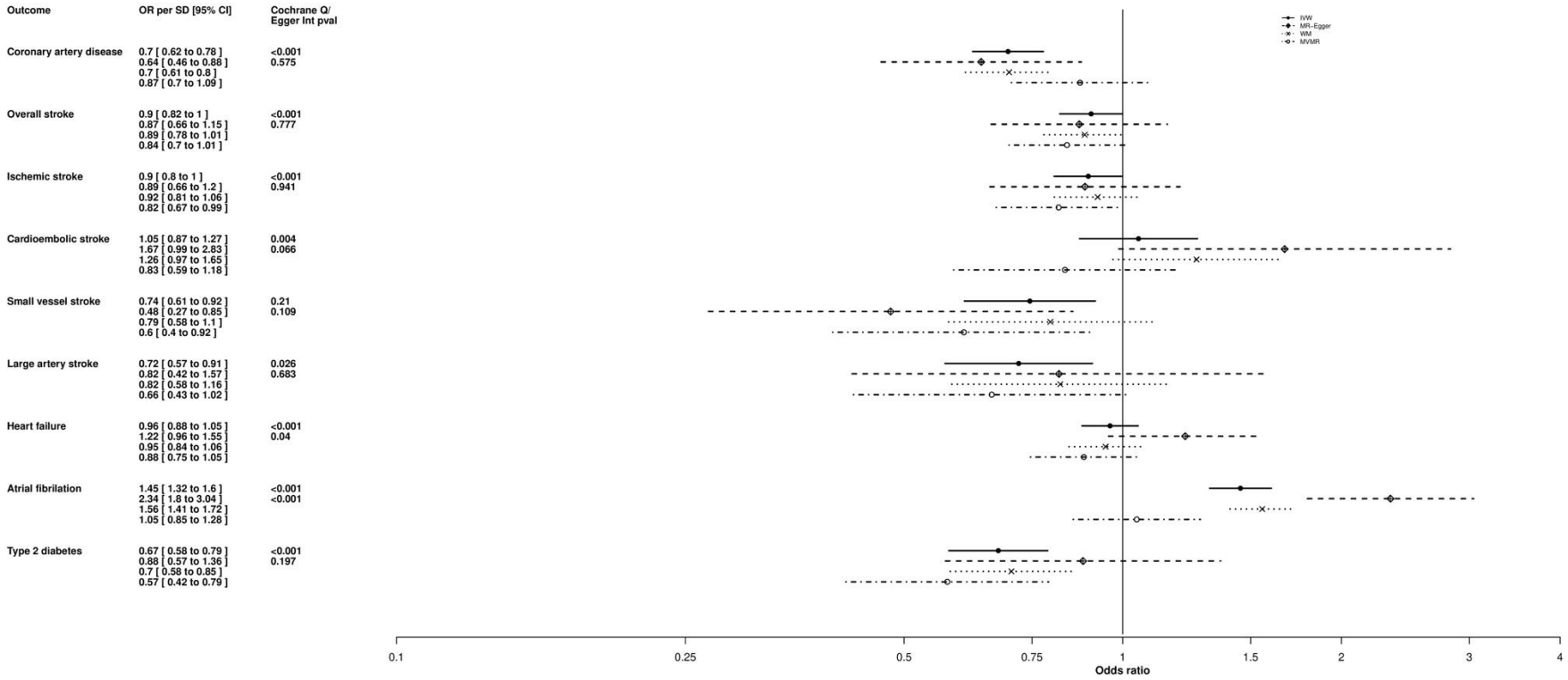
Exposure	Data source (PMID)	Sample size (% cases)	% European
Forced expiratory volume in 1 second (SD)	UK Biobank (IEU GWAS)	421,986	100
Forced vital capacity (SD)	UK Biobank (IEU GWAS)	421,986	100
Outcomes	**Data source (PMID)**	**Sample size (% cases)**	**% European**
Coronary artery disease	CARDIoGRAM (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 (FEV_1, per SD) on cardiovascular disease and type 2 diabetes using Mendelian randomization



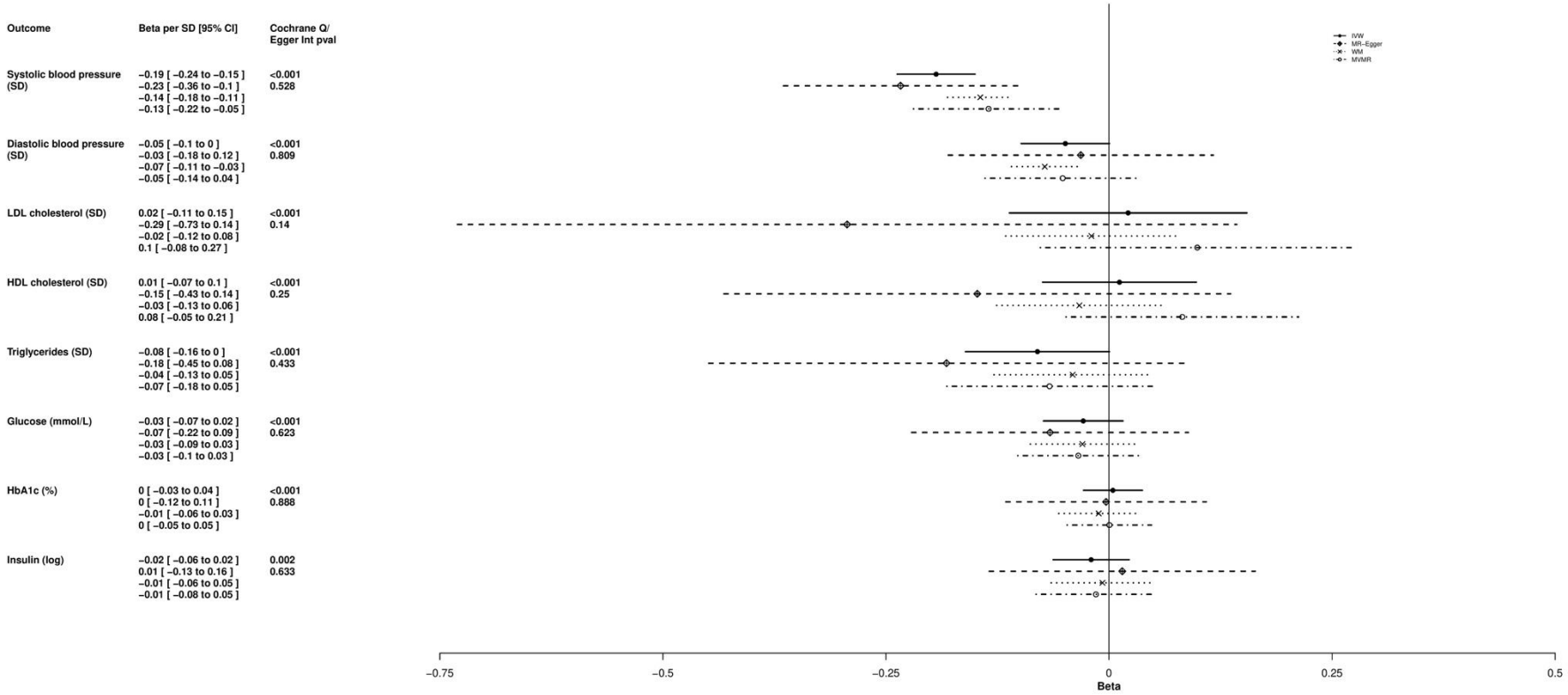
Footnote: *IVW: Inverse variance weighting; WM: Weighted median; MVMR: Multivariable Mendelian randomization (Adjusted for height)

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



Footnote: *IVW: Inverse variance weighting; WM: Weighted median; MVMR: Multivariable Mendelian randomization (Adjusted for height)

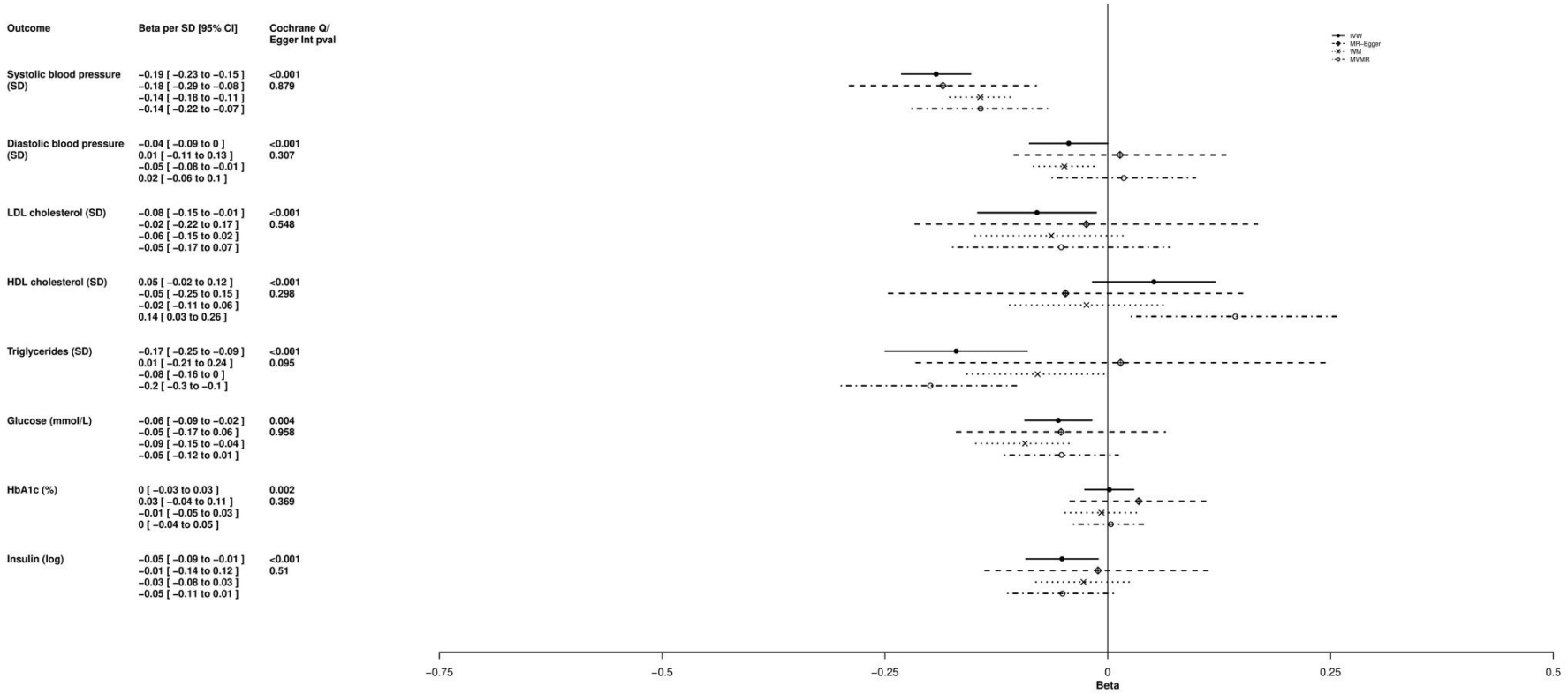
Figure 3: The impact of forced expiratory volume in 1 second (FEV_1, per SD) on cardiovascular risk factors using



Mendelian randomization

Footnote: *IVW: Inverse variance weighting; WM: Weighted median; MVMR: Multivariable Mendelian randomization (Adjusted for height)

Figure 4: The impact of forced vital capacity (FVC, per SD) on cardiovascular risk factors using Mendelian randomization



Footnote: *IVW: Inverse variance weighting; WM: Weighted median; MVMR: Multivariable Mendelian randomization (Adjusted for height)

Figure 5: The impact of higher predisposition to cardiovascular outcome on forced expiratory volume in 1 second (FEV_1) and forced vital capacity (FVC) using Mendelian randomization