ORIGINAL RESEARCH

Mental Health as a Mediator of the Association Between Educational Inequality and Cardiovascular Disease: A Mendelian Randomization Study

Daniel P. Jones [iD], BMBCh; Robyn E. Wootton, PhD; Dipender Gill [iD], BMBCh, PhD; Alice R. Carter [iD], PhD; David Gunnell, MD, PhD; Marcus R. Munafò, PhD; Hannah M. Sallis, PhD

BACKGROUND: Education is inversely associated with cardiovascular disease (CVD). Several mediators of this have been established; however, a proportion of the protective effect remains unaccounted for. Mental health is a proposed mediator, but current evidence is mixed and subject to bias from confounding factors and reverse causation. Mendelian randomization is an instrumental variable technique that uses genetic proxies for exposures and mediators to reduce such bias.

METHODS AND RESULTS: We performed logistic regression and 2-step Mendelian randomization analyses using UK Biobank data and genetic summary statistics to investigate whether educational attainment affects risk of mental health disorders. We then performed mediation analyses to explore whether mental health disorders mediate the association between educational attainment and cardiovascular risk. Higher levels of educational attainment were associated with reduced depression, anxiety, and CVD in observational analyses (odds ratio [OR], 0.79 [95% CI, 0.77–0.81], 0.76 [95% CI, 0.73–0.79], and 0.75 [95% CI, 0.74–0.76], respectively), and Mendelian randomization analyses provided evidence of causality (OR, 0.72 [95% CI, 0.67–0.77], 0.50 [95% CI, 0.42–0.59], and 0.62 [95% CI, 0.58–0.66], respectively). Both anxiety and depression were associated with CVD in observational analyses (OR, 1.63 [95% CI, 1.49–1.79] and 1.70 [95% CI, 1.59–1.82], respectively) but only depression showed evidence of causality in the Mendelian randomization analyses (OR, 1.09; 95% CI, 1.03–1.15). An estimated 2% of the total protective effect of education on CVD was mediated by depression.

CONCLUSIONS: Higher levels of educational attainment protect against mental health disorders, and reduced depression accounts for a small proportion of the total protective effect of education on CVD.

Key Words: anxiety ■ cardiovascular disease ■ depression ■ education ■ Mendelian randomization

See Editorial by Williams

Cardiovascular disease (CVD) is a leading cause of global morbidity and mortality.[1] The association between socioeconomic inequality and CVD is well-established, with individuals living in deprived areas typically having much higher cardiovascular mortality than those in less deprived areas.[2] More specifically, disparities in educational attainment have recently been shown to affect CVD using both conventional and genetic epidemiological techniques.[3] CVD risk was estimated to decrease by a third for every 3.6 years of additional full-time education past the age of 11 years.[4]

Several well-recognized cardiovascular risk factors appear to act as mediators for education.[5] For example, low educational attainment is associated with increased tobacco smoking, higher body mass index, and higher blood pressure, which, in turn, are linked to increased risk of CVD.[6] However, after accounting for

CLINICAL PERSPECTIVE

What Is New?

- We report a negative association between educational attainment and mental health conditions, with Mendelian randomization methods providing evidence for a causal effect.
- We report a positive association between depression and cardiovascular disease with Mendelian randomization methods providing evidence for a causal effect.
- Using conventional observational epidemiology and Mendelian randomization, we estimate that 2% of the protective effect of education on cardiovascular disease is mediated by depression; a sizeable proportion of this mediating effect by depression could be caused by tobacco smoking.

What Are the Clinical Implications?

- This study provides further evidence for the complex interplay of physical health, mental health, and social determinants regarding cardiovascular disease.
- This study also provides further evidence for the importance of addressing smoking behavior in the context of clinical depression.
- Investment in educational attainment is likely to be beneficial for both mental health and cardiovascular health; clinical interventions that target mental health could be beneficial for cardiovascular health, especially for individuals in lower socioeconomic positions.

Nonstandard Abbreviations and Acronyms

CARDIoGRAMplusC4D	Coronary Artery Disease Genome-Wide Replication and Meta-Analysis (CARDIoGRAM) Plus the Coronary Artery Disease (C4D) Genetics
IVW	inverse-variance weighted
MR	Mendelian randomization

these mediators, there is a significant proportion of the association between education and CVD that remains unaccounted for.

Existing evidence suggests an association between mental health and CVD as well as an association between education and mental health.[7,8] Consequently, mental health may be another potential mediator between education and CVD. For example, educational inequality may lead to disparities in psychological development by affecting exposure to stressors, resources to cope with stress, and negative external societal perception.[9] Poor mental health could, in turn, lead to increased CVD risk with proposed mechanisms including chronic inflammation,[10] increased sympathetic nervous system activation,[11] and increased cortisol exposure.[12]

Conventional observational studies that have explored education, mental health, and CVD have, so far, proved inconclusive.[13,14] Furthermore, observational evidence can be subject to bias from confounding and reverse causation, and, specifically in the context of mediation studies, to measurement error and collider bias.[15]

Mendelian randomization (MR) is an instrumental variable technique that uses genetic variants as a proxy for an exposure of interest.[16] These variants are randomized at conception and therefore largely inherited independently from other variants affecting confounding factors. These variants are also unchanged throughout the lifetime and so are unlikely to be affected by outcomes of interest, therefore reducing bias from reverse causation. Thus, MR reduces the sources of bias that limit causal inference in conventional observational research and is an effective, complementary tool in investigating potential causality.

The core assumptions of MR are that the genetic variants must be associated with the exposure under investigation, must only affect the outcome via that exposure, and must not share any common causes with the outcome. Provided that these assumptions are met, then an average causal effect can be estimated for genetic compliers in the sample, ie, patients whose exposure status reliably depends on particular genetic variants. The direction of these effect estimates can then be used as evidence against a causal null hypothesis.

Here, we specifically used 2-step MR, employing summary data for each step. Two-step MR estimates the exposure-mediator, exposure-outcome, and mediator-outcome effects separately and is particularly useful when mediators are binary. We also utilized separate samples for the instrument and outcome summary data, thus providing greater sample sizes in some cases.

In this study, we used a complementary approach utilizing both observational and MR techniques to explore the relationship between education, mental health, and CVD. Elucidation of this causal pathway could have important implications for policymakers interested in both physical and mental health improvement.

METHODS

Data Availability

The data used for the observational analyses conducted in this study are available to qualified researchers upon application to the UK Biobank (https://www.ukbiobank.ac.uk/researchers/). The data used for the MR analyses conducted in this study are publicly available via the cited references. The scripts used to conduct the analyses are available at https://github.com/MRCIEU/education_mentalhealth_CVD/

Observational Analyses

Sample

Our observational analyses used data from the UK Biobank, a national health resource that recruited participants aged 39 to 72 years from 22 centers across the UK between 2006 and 2010. Details of recruitment are publicly available.[17] We excluded 814 individuals based on consent withdrawal, reported aneuploidy, or mismatch between reported and chromosomal sex. We then applied the Medical Research Council Integrative Epidemiology Unit quality control procedure to restrict the sample to individuals of European ancestry.[18] The final sample size with complete phenotype measures was 333 525, with a mean age of 56.9 years (SD=8) and 54% women.

Measures

Educational Attainment

Participants in the UK Biobank reported the levels of qualifications attained and the highest level of qualification was converted into number of years in education using the International Standard Classification of Education coding of educational attainment (Table S1). Mean educational attainment was 13.9 years (SD=5.1). For regression analyses, educational attainment was categorized according to qualification level (minimal/none=0; further education=1; higher education=2) and thus odds ratios (ORs) should be interpreted as per increase in qualification level (approximately 3–5 years of additional attainment).

Mental Health Problems

Individuals were classified as having had a mental health problem if they reported "Yes" to either of the routine survey questions: "Have you ever seen a general practitioner (GP) for nerves, anxiety, tension, or depression?" or "Have you ever seen a psychiatrist for nerves, anxiety, tension, or depression?"

Depression

Individuals were classified as having major depressive disorder if their hospital episode statistics and/or general practice linked health record showed a diagnosis of a depressive episode or recurrent episodes according to the *International Classification of Diseases, Tenth Revision* (*ICD-10*) codes F32 and F33.

Anxiety

Individuals were classified as having anxiety disorder if their health record showed such a diagnosis according to *International Classification of Diseases* (*ICD*) code F41.

The prevalence of these mediators in our sample was 34% for mental health problems, 2.9% for depression, and 1.4% for anxiety.

Cardiovascular Disease

Individuals were classified as having CVD if their health records showed at least 1 diagnosis of angina or myocardial infarction before recruitment classified by *ICD* codes I20 and I21, respectively. The prevalence of CVD in the sample at the time of recruitment was 6%.

Additional covariates included age, sex, smoking status, hypertension, obesity, and socioeconomic position. The methods for their classification and the respective prevalence figures are reported in Data S1 and Table S2.

MR Analyses

Genetic Instruments

Educational Attainment

A total of 1271 independent genome-wide significant loci ($P<5\times10^{-8}$) were identified in a sample of 1 131 881 individuals in a genome-wide association study (GWAS) meta-analysis of educational attainment by Lee and colleagues.[19] Educational attainment was measured as years in education based on highest reported qualification level. The GWAS meta-analyzed 71 cohorts of European ancestry measuring educational attainment. The median effect size of each allele of these lead variants was 1.7 weeks of schooling.

Mental Health Problems

Fourteen independent genome-wide significant loci were identified in a recent GWAS of a phenotype described as "broad depression" in UK Biobank data (113 769 cases and 208 811).[20] Here, we report this phenotype as "Having had a mental health problem" instead. Individuals were classified as having had a mental health problem if they reported "Yes" to either of the routine survey questions: "Have you ever seen a general practitioner (GP) for nerves, anxiety, tension, or depression?" or "Have you ever seen a psychiatrist for

nerves, anxiety, tension, or depression?" Patients with diagnosed bipolar disorder or schizophrenia or who were taking antipsychotic medications were excluded. A number of participants were also excluded by inter-relatedness filtering.

Depression

A total of 44 genome-wide significant loci were identified in a GWAS meta-analysis of major depressive disorder (135 458 cases and 344 901 controls) by Wray and colleagues.[21] A total of 29 samples containing individuals of European ancestry were included in the analysis. Major depressive disorder was defined by meeting *Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)* (*DSM-IV*), or *International Classification of Diseases, Ninth Revision/ICD-10* criteria either by clinician review of medical records or by expert interview.

Anxiety

A recent GWAS of anxiety disorder (7016 cases and 14 745 controls) identified 1 independent genome-wide significant locus.[22] Anxiety cases were recruited from 7 different cohorts of European ancestry participating in the Anxiety NeuroGenetics Study with cases confirmed according to *DSM-IV* criteria. To improve instrument power, we relaxed the significance threshold to $P<5\times10^{-5}$ and identified 108 independent significant loci following linkage disequilibrium clumping ($r^2<0.001$, distance >10 000 kb).

Cardiovascular Disease

We used summary data from a recent GWAS meta-analysis from the CARDIoGRAMplusC4D (Coronary Artery Disease Genome-Wide Replication and Meta-Analysis [CARDIoGRAM] Plus the Coronary Artery Disease [C4D] Genetics) consortium.[23] This included 48 cohorts with a total of 60 801 cases and 123 504 controls. The cohort studies typically defined a case as having a diagnosis of myocardial infarction or angina with several also requiring confirmation from angiographic evidence.

Statistical Analysis

All analyses were performed using R (version 3.6.2).[24] A directed acyclic graph for the proposed causal relationships investigated is shown in Figure 1.

Observational Analysis

First, we used logistic regression to investigate the association of educational attainment with CVD and with each of our mediators having adjusted for the effects of age and sex. Second, we analyzed the effect of each of our mediators on CVD and, in order to effectively isolate the mediator effects, adjusted for age, sex, educational
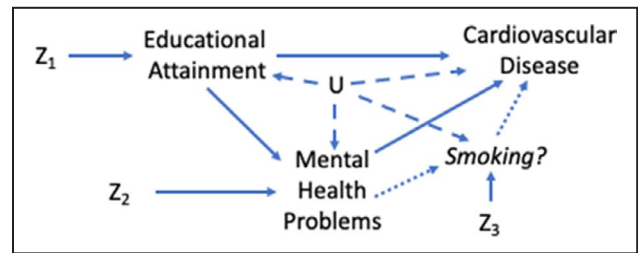

Figure 1. **A diagram illustrating the proposed causal relationships between education, mental health, and cardiovascular disease.**
Z_1, Z_2 and Z_3 represent genetic proxies and U represents confounding factors.

attainment, smoking status, socioeconomic position, hypertension, and obesity. When specifically analyzing the effects of depression on CVD, we additionally adjusted for the effects of anxiety and vice versa.

A proportion-mediated statistic was calculated using the product of coefficients method as detailed in previous work.[25] The effect sizes were calculated using absolute risk differences as estimated from the ORs from our regression analysis and the known prevalence figures in UK Biobank. Standard errors were calculated using the delta method. *P* values were calculated using *Z* tests.

MR Univariable Analysis

We conducted 2-sample MR analysis using the TwoSampleMR package 0.4.18.[26] We first ran our analysis using educational attainment as an exposure and each of our proposed mediators and CVD as outcomes. We then repeated the analysis using each mediator as an exposure and CVD as an outcome.

For these analyses, we used an inverse-variance weighted (IVW) approach as the main analysis,[27] with additional statistical sensitivity analyses: weighted mode,[28] weighted median,[29] and MR-Egger.[30] Each sensitivity analysis makes different assumptions about pleiotropy, therefore a consistent estimate among all 4 methods provides the strongest evidence. Effect estimates produced were interpreted for directionality and, thus, whether they provided evidence against a causal null hypothesis.

When using the anxiety disorder instrument, we performed MR robust adjusted profile score in addition to IVW to improve estimation while using a relaxed *P* value threshold for the instrument.[31] MR robust adjusted profile score offers a greater robustness to pleiotropy when using many weak instruments.

Two-Step MR Analysis

We next performed 2-step MR using the TwoSampleMR package 0.4.18. We used the results of our univariable

MR analysis for the exposure-mediator step. For the mediator-outcome step we then used the residual multivariable MR method to calculate the effect of each mediator on CVD having adjusted for the effects of educational attainment.[32] The estimates for each step were multiplied together to produce an indirect effect estimate and the proportion-mediated statistic was then calculated as for the observational analysis (detailed above).

Sensitivity Analyses

The methods for our sensitivity analyses are presented in Data S1. These analyses included MR-Egger, Cochran Q Tests of Heterogeneity, mean of F statistics, MR Steiger, and bidirectional observational models. They also included an observational analysis of mental health and CVD conducted without adjustment for other covariates except for age and sex. Sensitivity plots were also included for genetic instruments displaying susceptibility to bias from pleiotropy as well as MR analyses displaying significant heterogeneity.

Exploratory Analysis

To illustrate an approach for investigating multiple mediators between mental health and CVD, we performed an exploratory analysis of a likely candidate based on existing evidence. This analysis investigates the role of tobacco smoking as a mediator between depression and CVD. The methods are in Data S1.

Ethical Considerations

UK Biobank has received ethics approval from the UK National Health Service's National Research Ethics Service (reference 11/NW/0382).

RESULTS

Association of Education With CVD

We found strong evidence to suggest that educational attainment has an inverse association with CVD (Figure 2). In our observational analysis, increased educational attainment was associated with a reduction in CVD (OR, 0.75; 95% CI, 0.74–0.76 per qualification level increase in attainment). Our MR analysis provided evidence for a causal, protective effect of education on CVD (IVW OR, 0.62; 95% CI, 0.58–0.66).

Association of Education With Mental Health

We also found that education has an inverse association with anxiety and depression and that this is consistent among both adjusted observational and MR analyses (Figure 2). Increased educational attainment was associated with reductions in depression and anxiety (OR, 0.79 [95% CI, 0.77–0.81] and 0.76 [95% CI, 0.73–0.79], respectively, per qualification level increase in attainment). Our MR analysis provided evidence for education having a causal, protective effect against anxiety and depression (IVW OR, 0.50 [95% CI, 0.42–0.59] and 0.72 [95% CI, 0.67–0.77], respectively).

Educational attainment generally showed only weak evidence of an association with mental health problems in our observational analysis (OR, 0.99; 95% CI, 0.98–1.01 [$P=0.06$]) but, in our MR analysis, there was some evidence of a causal effect (IVW OR, 0.95; 95% CI, 0.94–0.96 [$P<0.0001$]). This association was consistent among the MR sensitivity analyses.

Association of Mental Health With CVD

In our observational analysis, mental health problems in general were associated with increased CVD risk

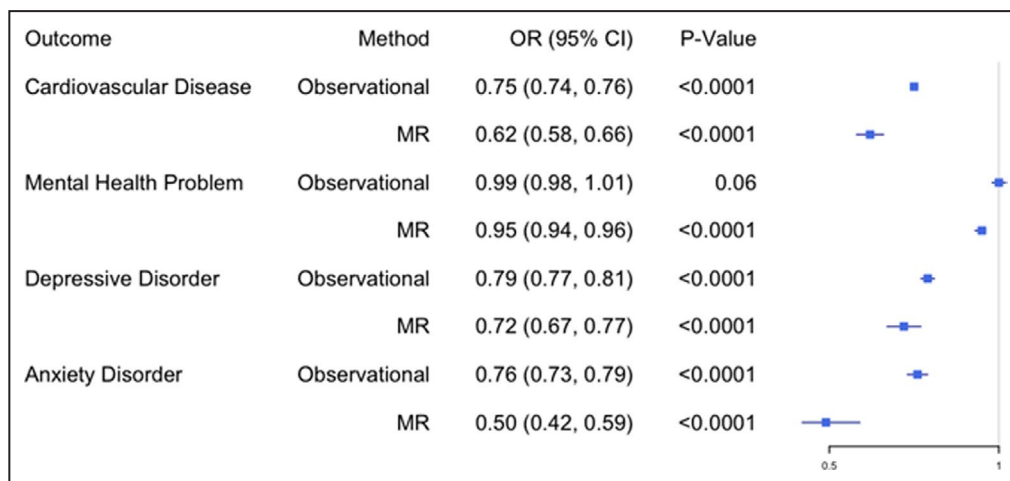| Outcome	Method	OR (95% CI)	P-Value
Cardiovascular Disease	Observational	0.75 (0.74, 0.76)	<0.0001
	MR	0.62 (0.58, 0.66)	<0.0001
Mental Health Problem	Observational	0.99 (0.98, 1.01)	0.06
	MR	0.95 (0.94, 0.96)	<0.0001
Depressive Disorder	Observational	0.79 (0.77, 0.81)	<0.0001
	MR	0.72 (0.67, 0.77)	<0.0001
Anxiety Disorder	Observational	0.76 (0.73, 0.79)	<0.0001
	MR	0.50 (0.42, 0.59)	<0.0001

Figure 2. The association of educational attainment with mental health and cardiovascular disease presented as odds ratios (ORs) per qualification level increase in attainment.
MR indicates Mendelian randomization.

(OR, 1.29; 95% CI, 1.24–1.33 [*P*<0.001]), but our MR analysis did not provide clear support for a causal interpretation (2-step MR: OR, 1.48; 95% CI, 0.92–2.39 [*P*=0.11]) (Figure 3).

However, we found strong observational evidence that depression was associated with increased cardiovascular risk (OR, 1.70; 95% CI, 1.58–1.82 [*P*<0.001]), and our MR analyses provided evidence for a causal effect (2-step MR: OR, 1.09; 95% CI, 1.03–1.15 [*P*=0.0021]). Anxiety disorder was strongly associated with CVD in our observational analysis (OR, 1.63; 95% CI, 1.49–1.79 [*P*<0.001]) but MR analysis did not support a causal interpretation (2-step MR: OR, 1.00; 95% CI, 0.99–1.01 [*P*=0.82]).

Proportion of the Association Between Education and CVD Mediated By Mental Health

Our mediation results show that mental health problems in general (observational: 0.3% [95% CI, 0.0%–2.8%]; 2-step MR: 4.2% [95% CI, 3.0%–5.3%]) and specifically depression (observational: 1.8% [95% CI, 1.2%–2.3%]; 2-step MR: 5.8% [95% CI, 4.8%–6.8%]) account for a small proportion of the negative association between education and CVD (Figures 4 and 5). The proportion mediated by anxiety disorder was small in the observational analysis, and we found little evidence of mediation in our MR analysis (observational: 0.9% [95% CI, 0.0%–0.6%]; 2-step MR: 0.0% [95% CI, 0.0%–0.6%]).

Sensitivity Analysis

The results of our sensitivity analyses are presented in Tables S3 through S13 and included testing for pleiotropy and reverse causation, heterogeneity, bidirectionality in our observational data, and personality traits as additional potential mediators. The results did raise the possibility of reverse causation between CVD and the proposed mediators in our observational analyses. There was also evidence of horizontal pleiotropy for the educational attainment and lifetime smoking instruments, which both displayed disparities between IVW and MR-Egger results. The MR-Egger intercept also showed weak evidence of directional horizontal pleiotropy for the lifetime smoking instrument (OR, 1.01; 95% CI, 1.00–1.02 [*P*=0.02]) although no other MR-Egger intercepts did. Finally, several of the analyses, particularly those utilizing educational attainment as an exposure, displayed significant effect heterogeneity.

Exploratory Analysis of Tobacco Smoking as a Mediator Between Depression and CVD

In exploratory analysis, we investigated smoking as a possible mediator on the pathway from depression to CVD. We found that depression is associated with increased smoking activity and that lifetime smoking was associated with increased CVD risk (Figures S1 and S2). In Figures 6 and 7, we show that smoking accounts for a large proportion of the increased CVD risk associated with depression (observational: 12.2% [95% CI, 11.6%–12.8%]; 2-step MR: 29% [95% CI, 28.9%–29.7%]).

DISCUSSION

In this study, we used both observational and MR techniques to investigate mental health as a mediator between education and CVD. Our results demonstrate that depression accounts for a proportion of the protective effect education has on CVD.

We demonstrated, in observational analyses, that educational attainment has a clear, inverse association with anxiety disorder, major depressive disorder, and CVD. Our MR analyses provided evidence for
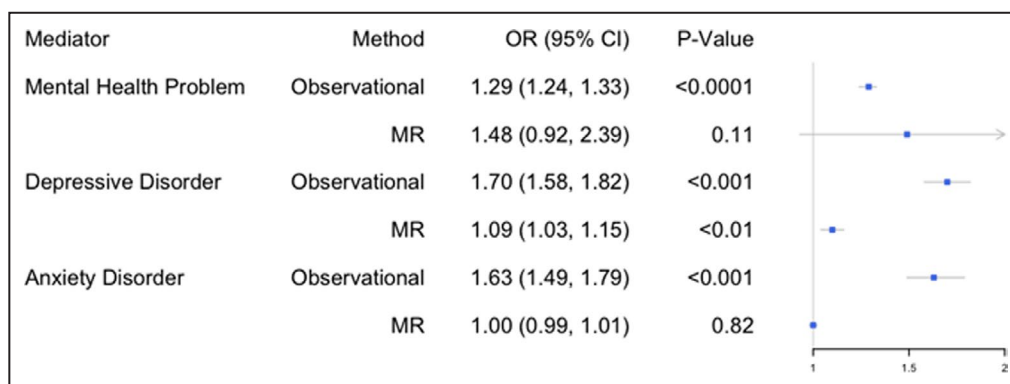
Mediator	Method	OR (95% CI)	P-Value
Mental Health Problem	Observational	1.29 (1.24, 1.33)	<0.0001
	MR	1.48 (0.92, 2.39)	0.11
Depressive Disorder	Observational	1.70 (1.58, 1.82)	<0.001
	MR	1.09 (1.03, 1.15)	<0.01
Anxiety Disorder	Observational	1.63 (1.49, 1.79)	<0.001
	MR	1.00 (0.99, 1.01)	0.82

Figure 3. **The association of mental health with cardiovascular disease following adjustment for the effects of educational attainment.**
OR indicates odds ratio; and MR, Mendelian randomization.

Figure 4. The estimated proportions of the inverse association between education and cardiovascular disease accounted for by mental health problems, depression, and anxiety.
MR indicates Mendelian randomization.

these relationships being causal. In addition, our MR analyses provided evidence for a causal, protective effect of educational attainment on mental health problems more broadly although observational analysis showed no strong association. Mental health problems in general and, specifically, anxiety and depression were all associated with increased cardiovascular risk in our observational analyses. Our MR analyses provided evidence only for causality between depression and CVD.

Following these initial results, we further explored the proposed causal relationships between education, mental health, and CVD, which likely consist of additional mediators (Figure 1). Our observational analyses showed that depression was positively associated with lifetime smoking, which, in turn, was associated with increased cardiovascular risk even after adjustment for

Figure 5. A diagram illustrating the breakdown of (A) the total effect of educational attainment on cardiovascular disease into (B) a direct effect and an indirect effect via depression.
OR indicates odds ratio (with 95% CI). Estimates are from Mendelian randomization (MR) analysis.

depression. Our MR analyses provided evidence for these associations being causal. We found that lifetime smoking accounts for as much as 30% of the association between depression and CVD.

Our Results in Context

Our analysis suggests that depression serves as a mediator between education and CVD, and thus accounts for an unexplained portion of the protective effect of education. Specifically, it suggests that educational disparities may produce additional CVD burden partly by increasing rates of depression. Previous MR analyses found no clear association between education and depression[33] and this was likely the result of weak instrument strength. Using stronger instruments in the current study, we did see novel evidence for a causal effect. In contrast, depression and coronary artery disease have previously been found to be associated using MR work,[34,35] and our mediation analysis constitutes a novel finding that can help identify possible new targets for interventions.

Our results support a protective role for education against mental health problems and, specifically, major depressive disorder and anxiety disorder. The foremost explanation for this is educational attainment being a proxy for overall socioeconomic position, which has a robust negative correlation with depression and anxiety.[36] For example, educational attainment is associated with both higher income and better occupation later in life.[37,38] Such factors protect against experiences such as debt, financial struggles, and poor working conditions, which feature among several relevant risk factors for mental health disorders.[39] Related interventions would accordingly include policies to promote educational attainment and reduce socioeconomic inequalities.

At an individual level, further years spent in education or training are also likely to play a protective role
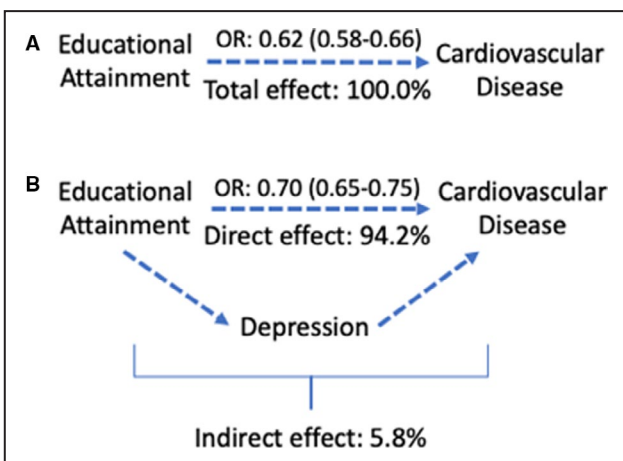
Mediator	Methods	% Proportion Mediated (95% CI)	P-value	
Smoking	Observational	12.2 (11.6, 12.8)	<0.0001	
	MR	29.3 (28.9, 29.7)	<0.0001	

0 2.5 5 7.5 10 12.5 15 17.5 20 22.5 25 27.5 30

Figure 6. The estimated proportion of the association between depression and cardiovascular disease accounted for by lifetime smoking.
MR indicates Mendelian randomization.

against mental health problems by directly modifying a range of psychosocial factors and providing resources that can be drawn on as stressors are encountered.[40] Such resources might include mental health advice, support networks, social participation, close friends, and good work-life balance.[41,42] Additional years in education may also help to offset exposure to stressors until one has greater maturity and resources to cope with them. By considering these observations, successful individual-level interventions in the education system might be developed.

In an exploratory analysis, we found that smoking accounts for a large proportion of the association between depression and CVD. The associations between smoking and CVD are well-known.[43,44] The association between smoking and depression has been commonly attributed to a "self-medication" hypothesis where smoking is used to cope with depressive symptoms. Other factors that are proposed to partially explain the remainder of the association between depression and CVD include chronic inflammation and changes in lipid metabolism.[45]

Figure 7. A diagram illustrating the breakdown of (A) the total effect of depression on cardiovascular disease into (B) a direct effect and an indirect effect via smoking.
OR indicates odds ratio (with 95% CI). Estimates are from Mendelian randomization analysis.

Implications

Our study has several important public health implications. First, it strengthens the case for investment in education interventions by demonstrating that educational attainment can reduce mental health burden, which is itself a significant source of morbidity and mortality.[46] Second, by illustrating the links between education, mental health, and CVD, we have reinforced the need to recognize the potential interplay between social determinants and physical and mental health problems. Indeed, our study shows how a physical disease burden could be partly alleviated by targeting a social determinant of mental health and provides further support for the use of existing mental health interventions to reduce risk for CVD.[47,48] Finally, we have also illustrated how to further explore potential mediators between depression and CVD in order to elucidate additional intervention targets. Here, we have highlighted tobacco smoking as such a target.

Strengths and Limitations

The strengths of our study include the use of MR analysis to complement a conventional observational approach, which allowed us to better overcome bias arising from confounding, measurement error, and reverse causation. The use of a "triangulation approach" also allowed for greater confidence in causal inference when there was consistency between the different approaches.[49] We were also able to limit bias arising from self-reporting and misclassification by using strict case definitions for both our observational and MR data sources. Although more relaxed case definitions may have improved the generalizability of results, potential bias arising from their use can be particularly problematic when using binary exposures in MR.[50] An additional strength of our study was the use of 2-step MR for a mediation analysis. This allowed us to specifically quantify the proportion of our exposure's effect attributable to our proposed mediators. Finally, we extended our analysis to illustrate how our methodology could explore multiple mediators in the proposed pathway.
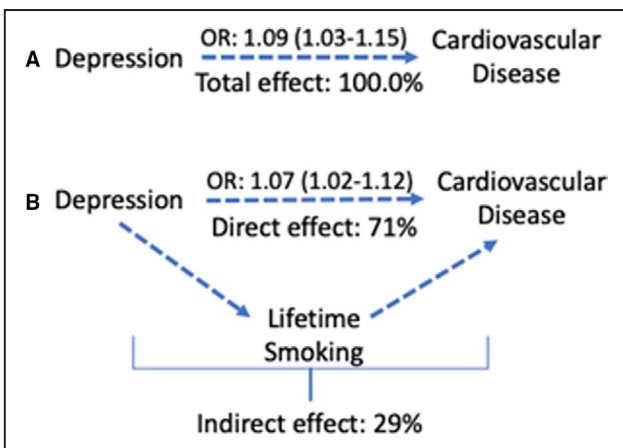
In our study, the use of MR demonstrated a discrepancy when observing the effects of mental health on CVD. When adjusting for confounding in our mediation analysis, we likely overadjusted in order to confidently isolate the effects of the mediator. To measure this, we conducted a sensitivity analysis without adjustment, although this suggested that the effect sizes we have found may actually be underestimates (Table S14). Instead, the discrepancy may have primarily been caused by reverse causation, with CVD leading to mental health problems and exaggerating observed associations. In a further sensitivity analysis, where CVD was the exposure and depression and then anxiety the outcomes, we did indeed find such bidirectionality (Table S15).

A common limitation of MR studies is weak instrument strength,[51] especially when strict case definitions are used. Our instruments for both educational attainment and depression were adequate in this regard, with large mean of F statistics and sample sizes. However, our instrument for anxiety had a relatively smaller mean of F statistic and required a relaxed genome-wide significance threshold; the null result should therefore be interpreted with caution and the analysis repeated when better MR instruments for anxiety are available. However, we utilized the MR robust adjusted profile score method to compensate for the relaxed threshold and we also conducted sensitivity analysis with the worry subscale of the Eysenck Personality Questionnaire (Revised Short Form). This analysis had adequate instrument strength and a larger sample size but also suggested no clear evidence (Table S8).

Despite adequate instrument strength, our "mental health problems" instrument produced wide CIs, suggesting that these analyses may have been underpowered. Unlike our other instruments, this one was constructed from a GWAS performed on a single sample, as the unusual phenotype measured is only available in the UK Biobank. The phenotype is also not clearly defined and is more heterogeneous than phenotypes using specific definitions for depression and anxiety. We believe this reiterates the need for future work to focus on generating instruments with larger sample sizes and specific phenotype definitions.

When conducting 2-sample MR, overlap between the samples can be problematic. The samples used for our exposure and mediators did feature partial overlap attributable to the use of UK Biobank data in both sets. This introduces a greater risk of a type I error in our 2-sample univariable analyses of education and mental health caused by possible overfitting. However, we also saw an inverse association with education in our sensitivity analyses that utilized personality traits related to mental health (Table S8). These featured substantively less overlap between samples. In addition, this limitation did not apply to our 2-step analysis, which pools the data for the exposure and mediators before analysis is conducted.

Another common limitation of MR studies is bias from directional pleiotropy.[52] We conducted numerous sensitivity analyses, including multiple MR methods and measurement of MR-Egger intercepts to assess for these. These suggested that the instruments for educational attainment and lifetime smoking could be biased by horizontal pleiotropy and that several related analyses also displayed effect heterogeneity. However, the direction of estimates was still largely consistent among a range of MR methods with sensitivity plots also demonstrating no specific variants driving these particular observations (Figures S3 through S12).

Despite this, caution should still be employed. Furthermore, the generalizability of our results is also limited given that our genetic instruments were solely constructed from cohorts of White European ancestry and that UK Biobank participants specifically skew to higher educational levels. This is in addition to another known issue with MR studies, namely that the genetic compliers producing the effect estimates may only constitute a subpopulation of the sample that may not be representative. Our instrument for educational attainment was based on a meta-analysis of several different populations, which would have helped to mitigate against these issues; however, our other genetic instruments were not as comprehensive. These limitations ultimately reiterate the need to interpret our MR results as evidence against a causal null hypothesis rather than as precise causal effect estimates. In the future, we hope that alternative genetic instruments based on a diverse range of subpopulations will be available to redress the issues highlighted.

Finally, even with our approach, education may still be acting as a proxy for socioeconomic position. A proposed strategy for investigating this would be to conduct an additional analysis using a measure such as household income, which would allow for comparison as well as adjustment as part of a multivariable MR approach. A recently published GWAS meta-analysis exploring household income[53] offers the possibility of this in future studies. Natural experiment techniques could also be used for this purpose as well as to support our results, having previously suggested that the raising of the school-leaving age reduced CVD.[54]

CONCLUSIONS

We have demonstrated that educational attainment is inversely associated with mental health problems, anxiety, depression, and CVD. A reduction in depression accounts for part of the inverse association between education and CVD. Finally, we have shown

that smoking accounts for a large proportion of the association between depression and CVD. Our findings offer significant clinical and public health implications by further demonstrating the interplay between social determinants, mental and physical disease, and associated targets for intervention.

ARTICLE INFORMATION

Affiliations

Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK (D.P.J., R.E.W., A.R.C., D.G., H.M.S.); Division of Population Medicine, School of Medicine, Cardiff University, UK (D.P.J.); MRC Integrative Epidemiology Unit, University of Bristol, Oakfield House, Bristol, UK (R.E.W., A.R.C., M.R.M., H.M.S.); Department of Epidemiology and Biostatistics, Imperial College London, London, UK (D.G.); Clinical Pharmacology and Therapeutics Section, Institute of Medical and Biomedical Education and Institute for Infection and Immunity, St George's, University of London, UK (D.G.); NIHR Biomedical Research Centre at the University Hospitals Bristol and Weston NHS Foundation Trust, University of Bristol, UK (D.G., M.R.M.); School of Psychological Science, University of Bristol, UK (M.R.M., H.M.S.); and Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of BristolUK, (H.M.S.).

Disclosures

D.G. is employed part-time by Novo Nordisk. The remaining authors have no disclosures to report.

Supplementary Material

Data S1: Supplementary Methods
Table S1–S15
Figure S1–S12
References 55–60

REFERENCES

1. Global atlas on cardiovascular disease prevention and control. WHO. https://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/. Published January 2020. Accessed January 25, 2020.
2. Inequalities in health. GOV.UK. https://www.gov.uk/government/publications/health-profile-for-england-2018/chapter-5-inequalities-in-health. Published 2018. Accessed January 25, 2020.
3. Rosengren A, Subramanian S, Islam S, Chow C, Avezum A, Kazmi K, Sliwa K, Zubaid M, Rangarajan S, Yusuf S. Education and risk for acute myocardial infarction in 52 high, middle and low-income countries: INTERHEART case-control study. *Heart*. 2009;95:2014–2022. DOI: 10.1136/hrt.2009.182436.
4. Tillmann T, Vaucher J, Okbay A, Pikhart H, Peasey A, Kubinova R, Pajak A, Tamosiunas A, Malyutina S, Hartwig FP, et al. Education and coronary heart disease: Mendelian randomisation study. *BMJ*. 2017;358:j3542. DOI: 10.1136/bmj.j3542.
5. Lynch J, Kaplan G, Cohen R, Tuomilehto J, Salonen J. Do cardiovascular risk factors explain the relation between socioeconomic status, risk of all-cause mortality, cardiovascular mortality, and acute myocardial infarction? *Am J Epidemiol*. 1996;144:934–942. DOI: 10.1093/oxfordjournals.aje.a008863.
6. Carter AR, Gill D, Davies NM, Taylor AE, Tillmann T, Vaucher J, Wootton RE, Munafò MR, Hemani G, Malik R, et al. Understanding the consequences of education inequality on cardiovascular disease: Mendelian randomisation study. *BMJ*. 2019;365:l1855. DOI: 10.1136/bmj.l1855.
7. Hemingway H, Marmot M. Evidence based cardiology: psychosocial factors in the aetiology and prognosis of coronary heart disease: systematic review of prospective cohort studies. *BMJ*. 1999;318:1460–1467. DOI: 10.1136/bmj.318.7196.1460.
8. Melkevik O, Hauge L, Bendtsen P, Reneflot A, Mykletun A, Aarø L. Associations between delayed completion of high school and educational attainment and symptom levels of anxiety and depression in adulthood. *BMC Psychiatry*. 2016;16:64. DOI: 10.1186/s12888-016-0765-1.
9. Gallo L, Matthews K. Understanding the association between socioeconomic status and physical health: do negative emotions play a role? *Psychol Bull*. 2003;129:10–51. DOI: 10.1037/0033-2909.129.1.10.
10. Lazzarino A, Hamer M, Gaze D, Collinson P, Rumley A, Lowe G, Steptoe A. The interaction between systemic inflammation and psychosocial stress in the association with cardiac troponin elevation: a new approach to risk assessment and disease prevention. *Prev Med*. 2016;93:46–52. DOI: 10.1016/j.ypmed.2016.09.018.
11. Hering D, Lachowska K, Schlaich M. Role of the sympathetic nervous system in stress-mediated cardiovascular disease. *Curr Hypertens Rep*. 2015;17:80. DOI: 10.1007/s11906-015-0594-5.
12. Iob E, Steptoe A. Cardiovascular disease and hair cortisol: a novel biomarker of chronic stress. *Curr Cardiol Rep*. 2019;21:116. DOI: 10.1007/s11886-019-1208-7.
13. Matthews K, Gallo L, Taylor S. Are psychosocial factors mediators of socioeconomic status and health connections? *Ann N Y Acad Sci*. 2010;1186:146–173. DOI: 10.1111/j.1749-6632.2009.05332.x.
14. Boylan J, Cundiff J, Matthews K. Socioeconomic status and cardiovascular responses to standardized stressors. *Psychosom Med*. 2018;80:278–293. DOI: 10.1097/PSY.0000000000000561.
15. Carter AR, Sanderson E, Hammerton G, Richmond RC, Smith GD, Heron J, Taylor AE, Davies NM, Howe LD. Mendelian randomisation for mediation analysis: current methods and challenges for implementation. *Eur J Epidemiol*. 2021 May 7. Online ahead of print. DOI: 10.1007/s10654-021-00757-1.
16. Baiocchi M, Cheng J, Small D. Instrumental variable methods for causal inference. *Stat Med*. 2014;33:2297–2340. DOI: 10.1002/sim.6128.
17. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, 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:e1001779. DOI: 10.1371/journal.pmed.1001779.
18. Mitchell R, Hemani G, Dudding T, Corbin L, Harrison S & Paternoster L IEU quality control. Data.bris.ac.uk. https://data.bris.ac.uk/datasets/1ovaau5sxunp2cv8rcy88688v/UK%20Biobank%20Genetic%20Data_MRC%20IEU%20Quality%20Control%20version%202.pdf. Published 2019. Accessed February 24, 2020.
19. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, Nguyen-Viet TA, Bowers P, Sidorenko J, Karlsson Linnér R, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*. 2018;50:1112–1121. DOI: 10.1038/s41588-018-0147-3.
20. Howard D, Adams M, Shirali M, Clarke T, Marioni R, Davies G, Coleman J, Alloza C, Shen X, Barbu M, et al. Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat Commun*. 2018;9:1470. DOI: 10.1038/s41467-018-03819-3.
21. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, Adams MJ, Agerbo E, Air TM, Andlauer TM, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet*. 2018;50:668–681. DOI: 10.1038/s41588-018-0090-3.

22. Otowa T, Hek K, Lee M, Byrne E, Mirza S, Nivard M, Bigdeli T, Aggen S, Adkins D, Wolen A, et al. Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry*. 2016;21:1391–1399. DOI: 10.1038/mp.2015.197.

23. The CARDIoGRAMplusC4D Consortium. A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47:1121–1130. DOI: 10.1038/ng.3396.

24. R Core Team. R: the R Project for statistical computing. R-project.org. https://www.r-project.org/. Published 2014. Accessed January 26, 2020.

25. VanderWeele TJ. Mediation analysis: a practitioner's guide. *Ann Rev Public Health*. 2016;37:17–32. DOI: 10.1146/annurev-publhealth-032315-021402.

26. MR Base for two sample MR. MRCIEU. https://mrcieu.github.io/TwoSampleMR/. Published 2019. Accessed January 27, 2020.

27. Lawlor D, Harbord R, Sterne J, Timpson N, Davey SG. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27:1133–1163. DOI: 10.1002/sim.3034.

28. Hartwig F, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46:1985–1998. DOI: 10.1093/ije/dyx102.

29. Bowden J, Davey Smith G, Haycock P, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40:304–314. DOI: 10.1002/gepi.21965.

30. Burgess S, Thompson S. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32:377–389. DOI: 10.1007/s10654-017-0255-x.

31. Qingyuan Z, Jingshu W, Jack B, Dylan SS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Arvix*. Preprint posted online January 29, 2018. https://arxiv.org/abs/1801.09652

32. Burgess S, Daniel RM, Butterworth AS, Thompson SG; EPIC-InterAct Consortium. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. *Int J Epidemiol*. 2015;44:484–495. DOI: 10.1093/ije/dyu176.

33. Viinikainen J, Bryson A, Böckerman P, Elovainio M, Pitkänen N, Pulkki-Råback L, Lehtimäki T, Raitakari O, Pehkonen J. Does education protect against depression? Evidence from the Young Finns Study using Mendelian randomization. *Prev Med*. 2018;115:134–139. DOI: 10.1016/j.ypmed.2018.08.026.

34. Wu Q, Kling JM. Depression and the risk of myocardial infarction and coronary death: a meta-analysis of prospective cohort studies. *Medicine*. 2016;95:e2815. DOI: 10.1097/MD.0000000000002815.

35. Mulugeta A, Zhou A, King C, Hyppönen E. Association between major depressive disorder and multiple disease outcomes: a phenome-wide Mendelian randomisation study in the UK Biobank. *Mol Psychiatry*. 2019;25:1469–1476. DOI: 10.1038/s41380-019-0486-1.

36. Fryers T, Melzer D, Jenkins R. Social inequalities and the common mental disorders. *Soc Psychiatry Psychiatr Epidemiol*. 2003;38:229–237. DOI: 10.1007/s00127-003-0627-2.

37. Education and earnings. OECD. https://stats.oecd.org/Index.aspx?DataSetCode=EAG_EARNINGS. Published 2020. Accessed on 28 January 2020.

38. Andersen R, Van De Werfhorst H. Education and occupational status in 14 countries: the role of educational institutions and labour market coordination. *Br J Sociol*. 2010;61:336–355. DOI: 10.1111/j.1468-4446.2010.01315.x.

39. Weich S, Lewis G. Poverty, unemployment and the common mental disorders: a population- based cohort study. *BMJ*. 1998;317:115–119. DOI: 10.1136/bmj.317.7151.115.

40. Matthews KA, Gallo LC. Psychological perspectives on pathways linking socioeconomic status and physical health. *Annu Rev Psychol*. 2011;62:501–530. DOI: 10.1146/annurev.psych.031809.130711.

41. Aldabe B, Anderson R, Lyly-Yrjanainen M, Parent-Thirion A, Vermeylen G, Kelleher C, Niedhammer I. Contribution of material, occupational, and psychosocial factors in the explanation of social inequalities in health in 28 countries in Europe. *J Epidemiol Community Health*. 2010;65:1123–1131. DOI: 10.1136/jech.2009.102517.

42. Moor I, Rathmann K, Stronks K, Levin K, Spallek J, Richter M. Psychosocial and behavioural factors in the explanation of socio-economic inequalities in adolescent health: a multilevel analysis in 28 European and North American countries. *J Epidemiol Community Health*. 2014;68:912–921. DOI: 10.1136/jech-2014-203933.

43. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ*. 2004;328:1519. DOI: 10.1136/bmj.38142.554479.AE.

44. Taylor A, Fluharty M, Munafò M. Investigating the possible causal association between smoking and depression and anxiety using Mendelian randomisation meta-analysis: the CARTA consortium. *BMJ Open*. 2014;4:e006141. DOI: 10.1136/bmjopen-2014-006141.

45. Khandaker G, Zuber V, Rees J, Carvalho L, Mason A, Foley C, Gkatzionis A, Jones P, Burgess S. Shared mechanisms between coronary heart disease and depression: findings from a large UK general population-based cohort. *Mol Psychiatry*. 2019;25:1477–1486. DOI: 10.1038/s41380-019-0395-3.

46. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3:e442. DOI: 10.1371/journal.pmed.0030442.

47. Gulliksson M, Burell G, Vessby B, Lundin L, Toss H, Svärdsudd K. Randomized controlled trial of cognitive behavioral therapy vs standard treatment to prevent recurrent cardiovascular events in patients with coronary heart disease: Secondary Prevention in Uppsala Primary Health Care Project (SUPRIM). *Arch Intern Med*. 2011;171:134–140. DOI: 10.1001/archinternmed.2010.510.

48. Orth-Gomér K, Schneiderman N, Wang H, Walldin C, Blom M, Jernberg T. Stress reduction prolongs life in women with coronary disease. *Circ Cardiovasc Qual Outcomes*. 2009;2:25–32. DOI: 10.1161/CIRCOUTCOMES.108.812859.

49. Lawlor D, Tilling K, Davey-Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol*. 2016;45:1866–1886. DOI: 10.1093/ije/dyw314.

50. Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol*. 2018;33:947–952. DOI: 10.1007/s10654-018-0424-6.

51. Davies Neil M, Holmes Michael V, Davey-Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*. 2018;362:k60. DOI: 10.1136/bmj.k601.

52. Hemani G, Bowden J, Davey-Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet*. 2018;27:R195–R208. DOI: 10.1093/hmg/ddy163.

53. Hill WD, Davies NM, Ritchie SJ, Skene NG, Bryois J, Bell S, Di Angelantonio E, Roberts DJ, Xueyi S, Davies G, et al. Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nat Commun*. 2019;10:5741. DOI: 10.1038/s41467-019-13585-5.

54. Hamad R, Nguyen TT, Bhattacharya J, Glymour MM, Rehkopf DH. Educational attainment and cardiovascular disease in the United States: a quasi-experimental instrumental variables analysis. *PLoS Medicine*. 2019;16:e1002834. DOI: 10.1371/journal.pmed.1002834.

55. Bowden J, Del Greco MF, Minelli C, Davey-Smith G, Sheehan N, Thompson J. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. *Int J Epidemiol*. 2016;45:1961–1974. DOI: 10.1093/ije/dyw220.

56. Brion M, Shakhbazov K, Visscher P. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol*. 2012;42:1497–1501. DOI: 10.1093/ije/dyt179.

57. Greco MF, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med*. 2015;34:2926–2940. DOI: 10.1002/sim.6522.

58. Hemani G, Tilling K, Davey SG. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet*. 2017;13:e1007081. DOI: 10.1371/journal.pgen.1007081.

59. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, Savage JE, Hammerschlag AR, Skene NG, Muñoz-Manchado AB, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet*. 2018;50:920–927. DOI: 10.1038/s41588-018-0151-7.

60. Liu M, Jiang YU, Wedow R, Li Y, Brazel DM, Chen F, Datta G, Davila-Velderrain J, McGuire D, Tian C, 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;51:237–244. DOI: 10.1038/s41588-018-0307-5.

Supplemental Material

Data S1.

Supplemental Methods

Methods for co-variates used in observational analysis

For age and sex, individuals were classified according to reported age and sex at time of recruitment. Individual smoking status was classified according to whether an individual reported never smoking, being a previous smoker or being a current smoker. Individuals were classified as having hypertension if their health records showed a diagnosis as classified by ICD code I10 and obesity if their records showed a diagnosis as classified by ICD code E66. Socioeconomic position was estimated using Townsend Deprivation Index (TDI) at birth.

Methods for sensitivity analyses

Additional MR tests: The regression dilution I^2 statistic was used to assess the suitability of the genetic instrument for use in the MR Egger method[55]. If the I^2 statistic was less than 0.9, then SIMEX corrections were conducted. This statistic also serves to measure potential bias arising from measurement error with values close to 1 indicating minimal attenuation. We also assessed the strength of the genetic instruments using the Mean F-statistic, with a value above 10 indicating adequate instrument strength[56]. In addition, we conducted a Cochran's Q-test to assess for heterogeneity[57] and then utilised the MR Egger intercept in order to assess for directional pleiotropy. We also applied Steiger filtering directionality tests to assess for reverse causation by identifying genetic variants that explain more variance in the outcome than the exposure[58].

Bidirectionality: As a sensitivity analysis, we investigated the directionality of these effects by measuring the effect of cardiovascular disease on our proposed mediators having adjusted for age, sex and SES.

Additional mediators: We also investigated whether personality traits such as depressed effect and worry were mediating the relationship between educational attainment and cardiovascular disease, rather than mental health problems. These traits were depressed affect and worry with instruments created based on summary data from a recent GWAS meta-analysis of neuroticism[59]. This included 12 cohorts with a total of 449,484 individuals and identified 136 independent genome-wide significant loci for neuroticism, depressed affect and worry.

Smoking as a mediator between mental health problems and cardiovascular disease: Finally, we investigated whether smoking behaviours further mediated the pathway from mental health problems to cardiovascular disease using multivariable regression and MR approaches. For the observational analyses, we used UK Biobank data and classified participants on a never vs ever smoker basis. For the MR analyses, we constructed an instrument for smoking initiation using summary data from the GSCAN GWAS[60] which defined smoking initiation as ever having smoked more than 100 cigarettes, smoking every day for at least a month or smoking regularly. The sample size was 1,232,091 and it identified 378 independent genome-wide significant genetic loci which explain 4% of the variance in smoking initiation. We used the same summary data reporting cardiovascular disease for our outcome as before.

We utilised the product of co-efficients method as described in our main methods to calculate the proportion mediated statistic. Effects were estimated using logistic regression for our traditional observational data and two-step MR for our genetic instruments. In the first step when calculating the indirect effects, we adjusted for age, sex, educational attainment, obesity, socio-economic position and hypertension in our logistic regression. In the second step, we adjusted for age, sex and depression in the logistic regression and depression only in the two-step MR. Standard errors were calculated using the delta method.

Table S1. International Standard for Classification of Education (ISCED) codes applied to levels of qualification reported in UK biobank data.

Highest Level of Qualification	ISCED Code	Years of full-time Education
College/University Degree	5	20
National Vocational Qualification/equivalent	5	19
Other professional qualification	4	15
A-levels/AS/Equivalent	3	13
GCSE/CSE/O-Level/Equivalent	2	10
None of the above	1	7
Prefer not to say	N/A	Excluded

Table S2. Reported demographics in UK Biobank for co-variates utilised in observational analysis.

Co-Variate		Demographics
Age in years (mean (SD))		56.9 (8)
Sex (%)	Male	46
	Female	54
Smoking Status (%)	Never	55.2
	Former	35.5
	Current	9.3
Hypertension (%)		18.7
Obesity (%)		2.6
Townsend Deprivation Index (mean (SD))		-1.58 (2.93)

Table S3. Full results of MR univariable analysis.

Exposure	Outcome	Method	SNPs	OR (95% CI)	P-value
Educational Attainment	Cardiovascular Disease	Inverse Variance Weighted (IVW)	1206	0.62 (0.58-0.66)	3.5×10^{-50}
		MR Egger Simex (Unweighted)	1206	0.73 (0.58-0.91)	0.004
		Weighted Median	1206	0.64 (0.59-0.70)	1.1×10^{-24}
		Weighted Mode	1206	0.71 (0.47-1.06)	0.10
	Major depressive Disorder (MDD)	IVW	1068	0.72 (0.67-0.77)	8.0×10^{-22}
		MR Egger Simex (Unweighted)	1068	0.62 (0.51-0.75)	1.8×10^{-6}
		Weighted Median	1068	0.75 (0.69-0.81)	4.9×10^{-12}
		Weighted Mode	1068	0.98 (0.61-1.57)	0.94
	Anxiety Disorder	IVW	1151	0.50 (0.42-0.59)	8.8×10^{-16}
		MR Egger Simex (Weighted)	1151	0.49 (0.21-1.14)	0.10
		Weighted Median	1151	0.50 (0.39-0.66)	3.5×10^{-7}
		Weighted Mode	1151	0.60 (0.21-1.70)	0.33
	Mental Health Problems	IVW	568	0.95 (0.94-0.96)	7×10^{-12}
		MR Egger Simex (Unweighted)	568	0.97 (0.93-1.02)	0.20
		Weighted Median	568	0.95 (0.93-0.96)	7.2×10^{-11}
		Weighted Mode	568	0.91 (0.86-0.97)	0.002
MDD	Cardiovascular Disease	IVW	39	1.09 (0.98-1.22)	0.13
		Weighted Median	39	1.10 (0.95-1.28)	0.20
		Weighted Mode	39	1.20 (0.90-1.61)	0.23
Anxiety		IVW	94	1.01 (0.99-1.02)	0.27
		MR Egger Simex (Unweighted)	94	1.01 (0.98-1.04)	0.42
		Weighted Median	94	1.01 (0.99-1.03)	0.29
		Weighted Mode	94	1.01 (0.98-1.04)	0.59
		MR RAPS	94	1.01 (1.00-1.03)	0.13
Mental Health Problems		IVW	13	2.68 (1.32-5.44)	0.01
		Weighted Median	13	2.79 (1.06-7.33)	0.04
		Weighted Mode	13	2.41 (0.42-13.80)	0.34
MDD	Lifetime Smoking	IVW	35	1.07 (1.03-1.12)	1.5×10^{-3}
		Weighted Median	35	1.07 (1.04-1.10)	8.5×10^{-6}
		Weighted Mode	35	1.08 (1.02-1.14)	9.6×10^{-3}
Lifetime Smoking	Cardiovascular Disease	IVW	120	1.87 (1.49-2.36)	9.3×10^{-8}
		MR Egger Simex (Unweighted)	120	0.68 (0.31-1.53)	0.35
		Weighted Median	120	2.07 (1.53-2.79)	2.1×10^{-6}
		Weighted Mode	120	2.35 (1.00-5.52)	0.05

Table S4. Regression dilution I^2_{GX} to test the suitability of instruments for MR Egger and the mF statistic as a test of instrument strength for MR univariable analysis.

Exposure	Outcome	mF	I^2 Unweighted G_X	I^2 Weighted G_X
Educational Attainment	Cardiovascular disease	46.66	0.732	0.568
	Major Depressive Disorder (MDD)	46.64	0.736	0.556
	Anxiety	47.10	0.689	0.766
	Mental Health Presentation	48.37	0.770	0.574
MDD	Cardiovascular disease	36.63	0	0
Anxiety		18.51	0.524	0
Mental Health Problems		36.80	0	0
MDD	Lifetime Smoking	37.95	0	0
Lifetime Smoking	Cardiovascular disease	44.05	0.644	0.454

Table S5. The number of SNPs from the genetic instruments in the MR univariable analysis that explain more of the variance in the exposure than the outcome according to Steiger directionality tests.

Exposure	Outcome	True (%)	FALSE
Educational Attainment	Cardiovascular disease	1206 (95.3)	60
	Mental health problems	595 (100)	0
	Major Depressive Disorder (MDD)	1228 (96.6)	51
	Anxiety	702 (58.2)	505
Mental health problems	Cardiovascular disease	12 (92.3)	1
MDD		40 (100)	0
Anxiety		95 (100)	0
MDD	Lifetime Smoking	36 (100)	0
Lifetime Smoking	Cardiovascular Disease	116 (92.1)	10

Table S6. Cochran's Q Tests of Heterogeneity for MR univariable analysis.

Exposure	Outcome	IVW Q (Degrees of freedom)	IVQ Q P-value
Educational Attainment	Cardiovascular Disease	1499.43 (1205)	<0.001
	Mental Health Problems	1193.86 (567)	<0.001
	Major Depressive Disorder (MDD)	1901.98 (1067)	<0.001
	Anxiety	1111.07 (1150)	0.79
Mental Health Problems	Cardiovascular disease	11.72 (12)	0.47
MDD		47.81 (38)	0.13
Anxiety		95.85 (93)	0.40
MDD	Lifetime Smoking	261.97 (34)	<0.001
Lifetime Smoking	Cardiovascular disease	183.08 (119)	<0.001

Table S7. MR Egger Intercept test of directional horizontal pleiotropy for MR univariable analysis.

Exposure	Outcome	MR Egger Intercept	Standard Error	P-value
Educational Attainment	Cardiovascular Disease	-0.00144	0.00132	0.28
	Mental Health Problems	-0.000193	0.000286	0.50
	Major Depressive Disorder (MDD)	0.000999	0.00138	0.47
	Anxiety	-0.00197	0.00378	0.60
Mental Health Problems	Cardiovascular Disease	0.0193	0.0183	0.31
MDD		0.00729	0.00125	0.56
Anxiety		-0.000687	0.00293	0.82
MDD	Lifetime Smoking	0.00367	0.00479	0.45
Lifetime Smoking	Cardiovascular Disease	0.0122	0.00496	0.02

Table S8. Results of MR univariable analysis utilising personality traits.

Exposure	Outcome	Method	SNPs	OR (95% CI)	P-value
Educational Attainment	Depressed Affect	Inverse Variance Weighted (IVW)	1140	0.75 (0.73-0.77)	1.2×10^{-133}
		MR Egger	1140	0.74 (0.69-0.80)	5.9×10^{13}
		Weighted Median	1140	0.77 (0.59-0.70)	1.2×10^{-104}
		Weighted Mode	1140	0.76 (0.67-0.87)	6.3×10^{-5}
	Worry	IVW	1140	0.92 (0.90-0.94)	4.3×10^{-11}
		MR Egger	1140	0.96 (0.88-1.04)	0.32
		Weighted Median	1140	0.92 (0.89-0.94)	8.9×10^{-12}
		Weighted Mode	1140	0.87 (0.78-0.98)	0.02
Depressed Affect	Cardiovascular Disease	IVW	54	1.49 (1.18-1.87)	0.0007
		MR Egger	54	1.36 (0.41-4.50)	0.61
		Weighted Median	54	1.38 (1.07-1.77)	0.01
		Weighted Mode	54	1.41 (0.79-2.52)	0.26
Worry		IVW	54	0.95 (0.75-1.21)	0.68
		MR Egger	54	0.65 (0.18-2.36)	0.51
		Weighted Median	54	1.08 (0.83-1.41)	0.55
		Weighted Mode	54	1.14 (0.64-2.01)	0.66

Table S9. Results of multivariable MR analysis utilising personality traits with odds ratios for each mediator adjusted for effects of exposure presented.

Exposure	Mediator	Outcome	OR (95% CI)	P-value
Educational Attainment	Worry	Cardiovascular Disease	1.04 (0.97-1.11)	0.54
	Depressed Affect		1.10 (0.99-1.21)	0.06

Table S10. Cochran's Q Tests of Heterogeneity for multivariate MR analysis.

Exposure	Mediator	Outcome	Q (Degrees of freedom)	Q P-value
Educational Attainment	Major Depressive Disorder (MDD) Anxiety Mental Health Problems	Cardiovascular Disease	1534 (1269) 1677 (1271) 1782 (1329)	<0.001 <0.001 <0.001
Educational Attainment	MDD	Lifetime Smoking	3129 (1084)	<0.001
MDD	Lifetime Smoking	Cardiovascular Disease	189 (141)	<0.01

Table S11. MR Egger Intercept test of directional horizontal pleiotropy for MR multivariable analysis.

Exposure	Mediator	Outcome	MR Egger Intercept	Standard Error	P-value
Educational Attainment	Mental Health Problems	Cardiovascular Disease	0.00104	0.000532	0.05
	Major Depressive Disorder (MDD)		0.000555	0.000357	0.12
	Anxiety		0.000405	0.000354	0.25
Educational Attainment	MDD	Lifetime Smoking	0.0000204	0.0000851	0.81
MDD	Lifetime Smoking	Cardiovascular Disease	0.00236	0.00102	0.02

Table S12. The MR Egger mF statistic as a test of instrument strength in MR multivariable analysis.

Exposures	Mediator	Outcome	mF	P-value
Educational Attainment	Mental Health Problems	Cardiovascular disease	60.81	<0.001
	Major Depressive Disorder (MDD)		117.9	<0.001
	Anxiety		113.4	<0.001
Educational Attainment	MDD	Lifetime Smoking	533.9	<0.001
MDD	Lifetime Smoking	Cardiovascular disease	19.8	<0.001

Table S13. Results of MR univariable analysis of personality traits.

Exposure	Outcome	Method	SNPs	OR (95% CI)	P-value
Educational Attainment	Depressed Affect	Inverse Variance Weighted (IVW)	1140	0.75 (0.73-0.77)	1.2×10^{-133}
		MR Egger	1140	0.74 (0.69-0.80)	5.9×10^{13}
		Weighted Median	1140	0.77 (0.59-0.70)	1.2×10^{-104}
		Weighted Mode	1140	0.76 (0.67-0.87)	6.3×10^{-5}
Educational Attainment	Worry	IVW	1140	0.92 (0.90-0.94)	4.3×10^{-11}
		MR Egger	1140	0.96 (0.88-1.04)	0.32
		Weighted Median	1140	0.92 (0.89-0.94)	8.9×10^{-12}
		Weighted Mode	1140	0.87 (0.78-0.98)	0.02
Depressed Affect	Cardiovascular Disease	IVW	54	1.49 (1.18-1.87)	0.0007

Table S14. Results of logistic regression of mental health on cardiovascular disease with adjustment for age, sex and educational attainment only.

Mediator	Outcome	OR (95% CI)	P-value
Mental health problems	Cardiovascular disease	1.54 (1.48-1.60)	8.63×10^{-129}
Major depressive disorder		2.77 (2.60-2.96)	2.60×10^{-212}
Anxiety disorder		2.39 (2.19-2.61)	2.27×10^{-84}

\

Table S15. Results of logistic regression assessing the effect of cardiovascular disease on mental health utilising observational data after adjustment for co-variates.

Exposure	Outcome	OR (95% CI)	P-value
Cardiovascular Disease	Mental Health Problems	1.23 (1.19-1.26)	1.2×10^{-31}
	Anxiety Disorder	1.55 (1.40-1.72)	3.5×10^{-21}
	Major Depressive Disorder	1.57 (1.46-1.68)	9.1×10^{-38}

Figure S1. Results of univariable observational and MR analyses investigating association of depression with lifetime smoking index (LSI).

Exposure	Outcome	Method	Mean Difference (95% CI)	P-Value
Depression	LSI	Observational	+0.23 (0.22, 0.24)	<0.0001
		MR	+0.04 (0.03, 0.05)	<0.01

0 0.025 0.075 0.125 0.175 0.225
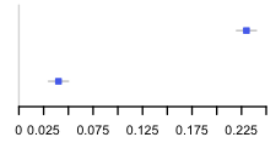
Figure S2. Results of multivariable observational and MR analyses investigating association of lifetime smoking with cardiovascular disease having adjusted for depression.

Mediator	Outcome	Method	OR (95% CI)	P-Value
LSI	CVD	Observational	1.32 (1.29, 1.34)	<0.0001
		MR	1.87 (1.49, 2.36)	<0.0001
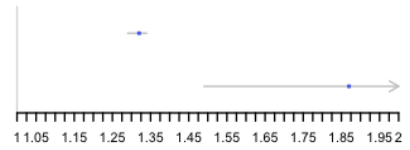
Figure S3. Plot of MR leave-one-out analysis of educational attainment and cardiovascular disease.

MR leave-one-out sensitivity analysis for
'Educational Attainment' on 'CVD'
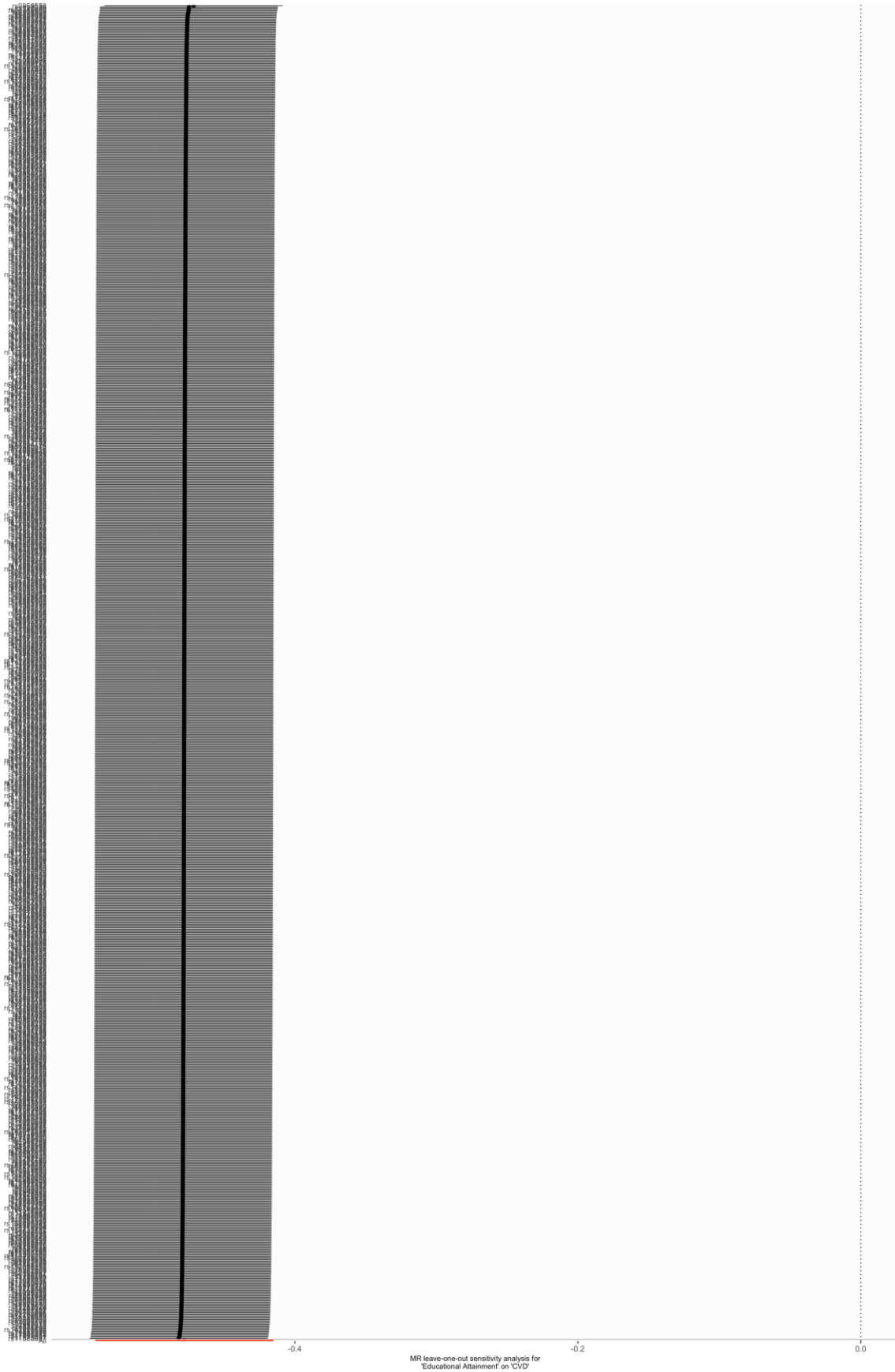
Figure S4. Scatter plot of MR analysis of educational attainment and cardiovascular disease.
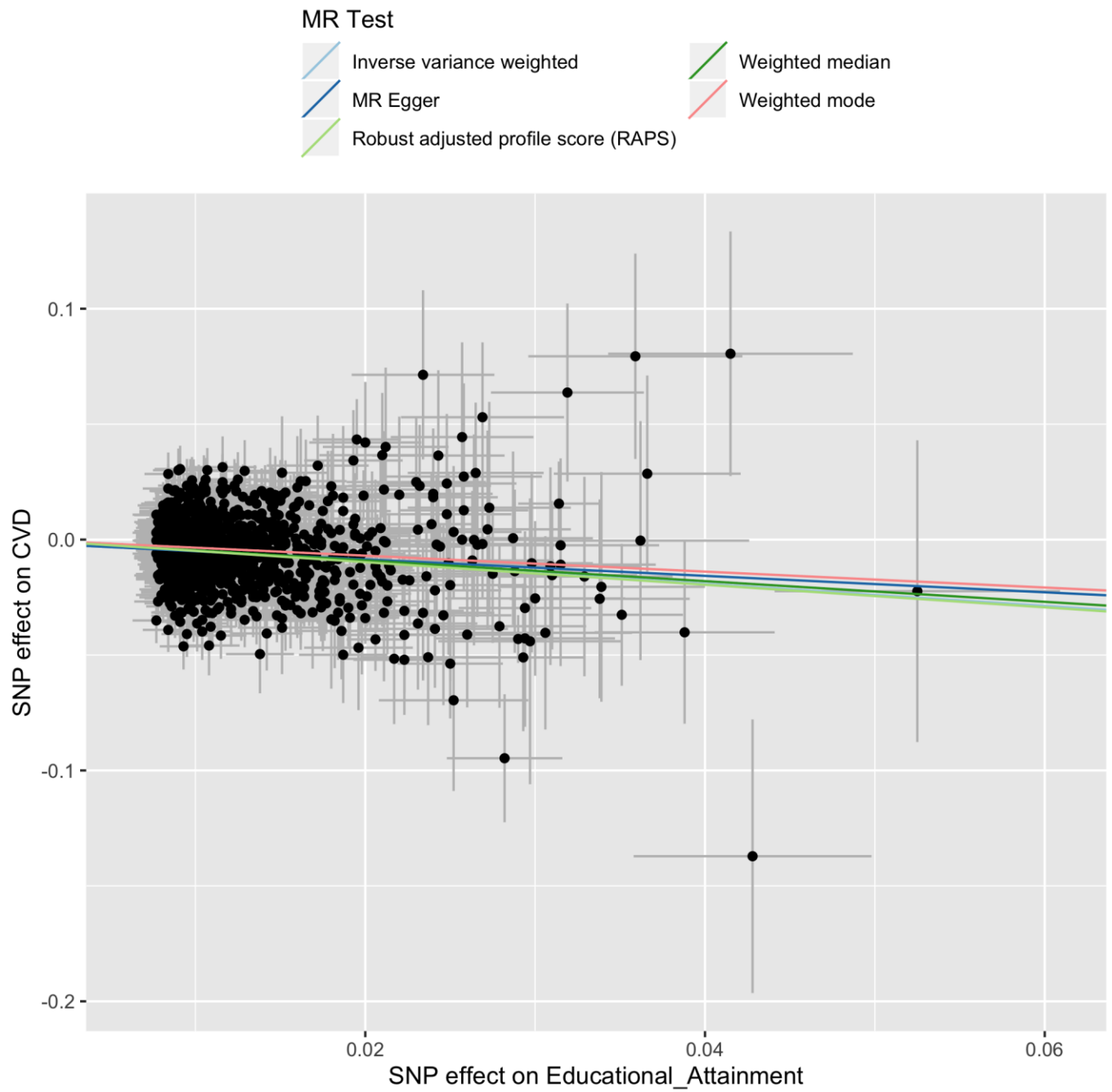
Figure S5. Plot of MR leave-one-out analysis of educational attainment and mental health problems.

MR leave-one-out sensitivity analysis for
'Educational Attainment' on 'Mental Health'
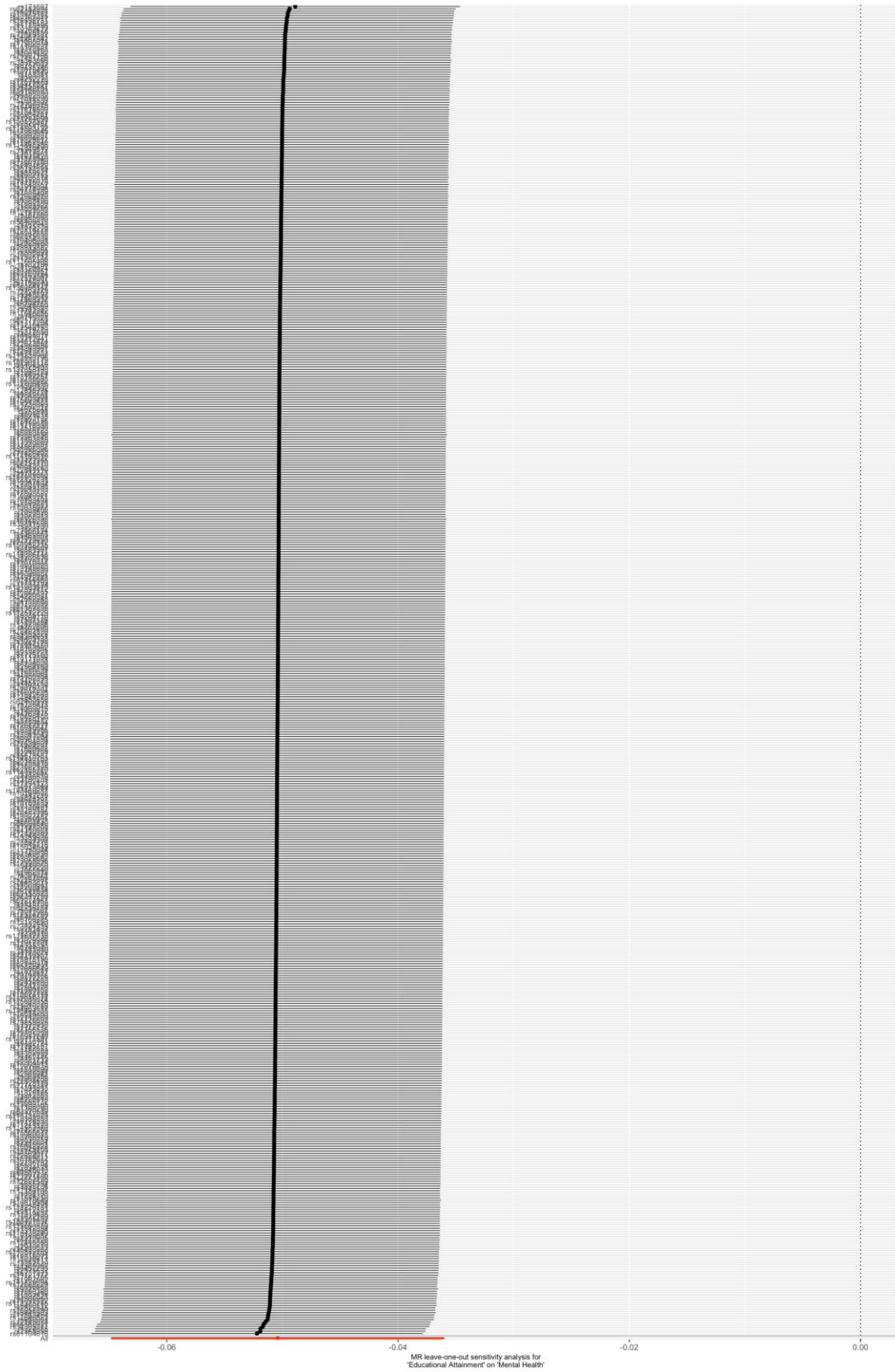
Figure S6. Scatter plot of MR analysis of educational attainment and mental health problems.
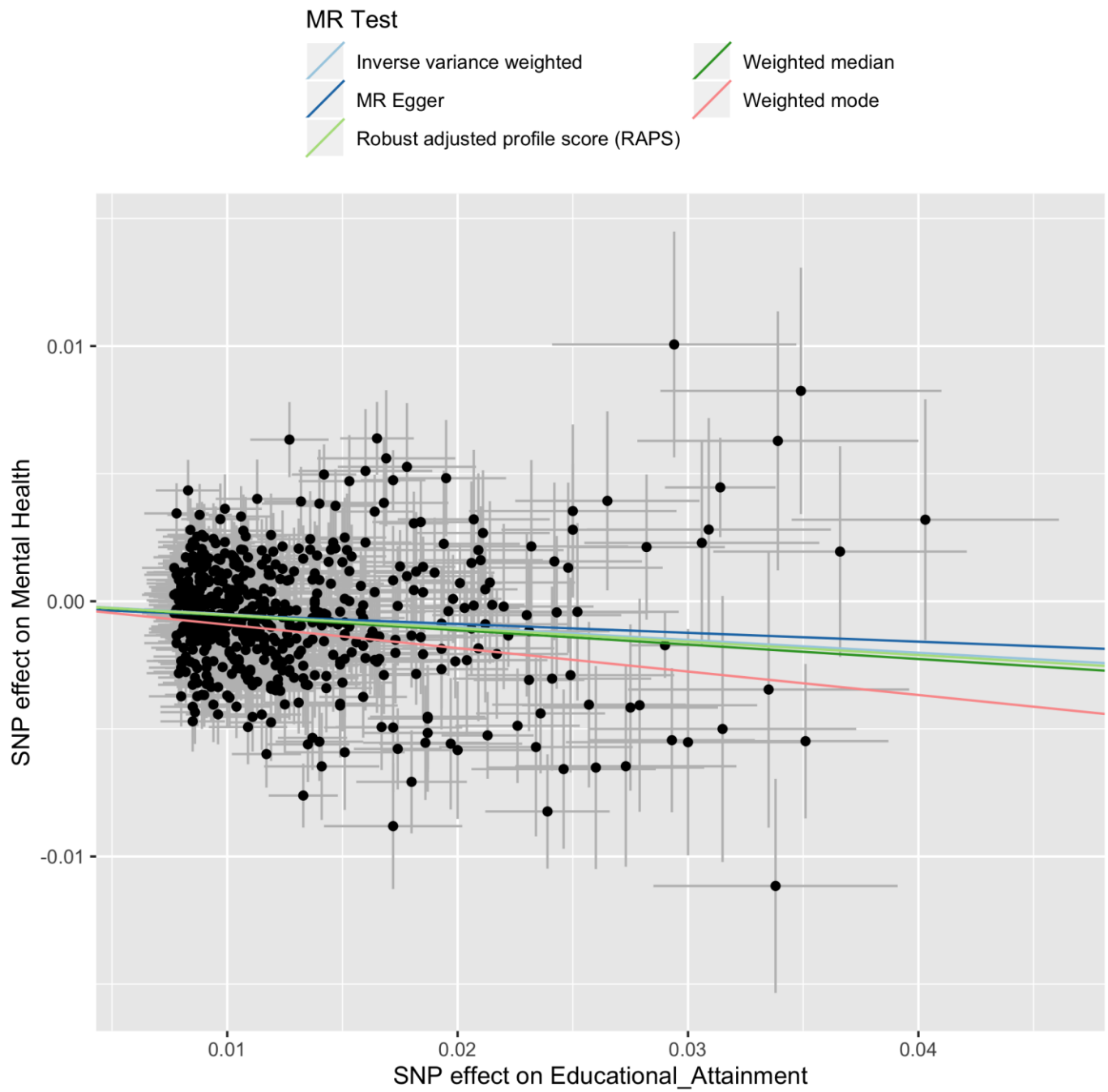
Figure S7. Plot of MR leave-one-out analysis of educational attainment and depression.

MR leave-one-out sensitivity analysis for
'Educational Attainment' on 'Depression'
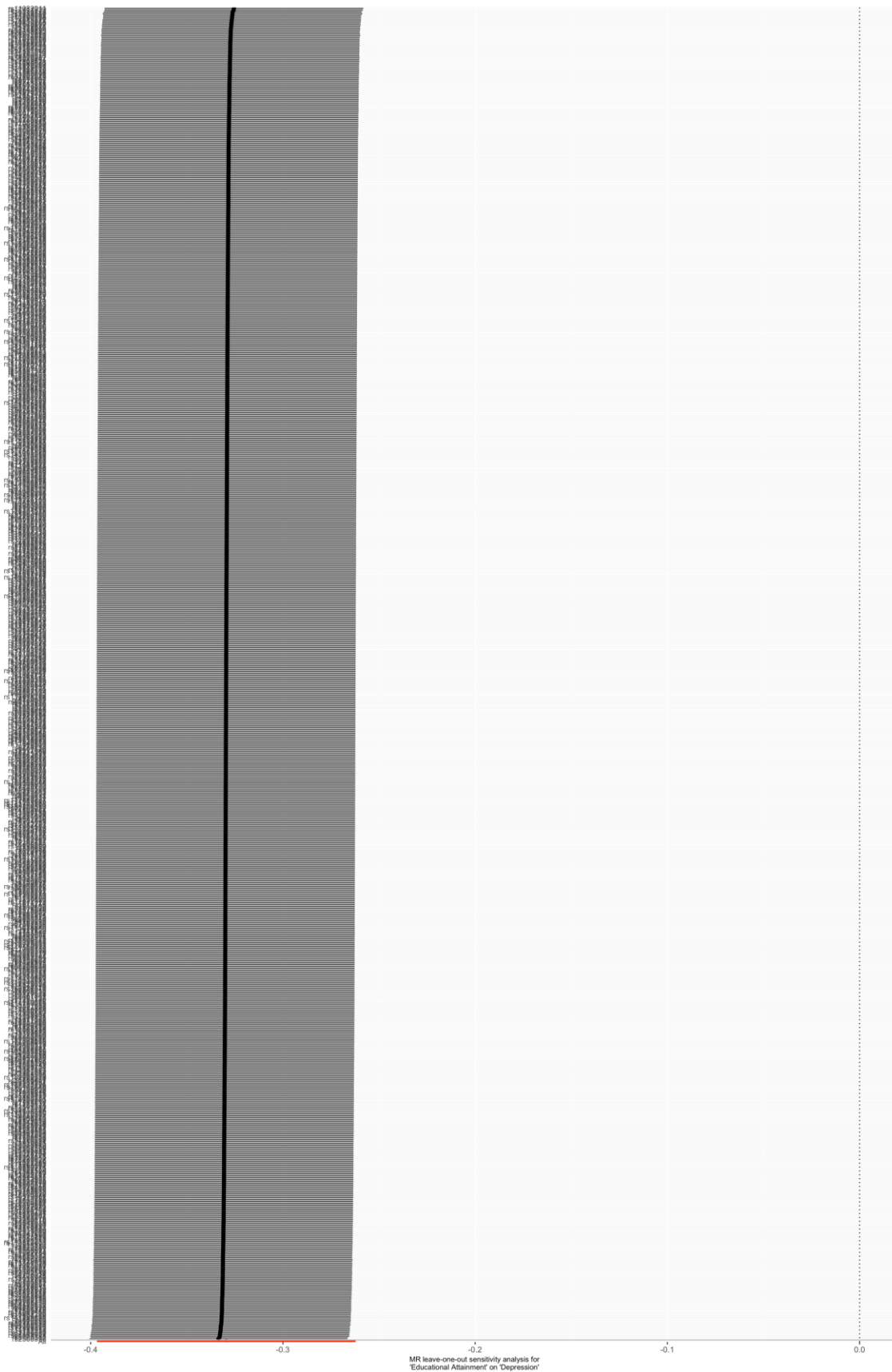
Figure S8. Scatter plot of MR analysis of educational attainment and depression.

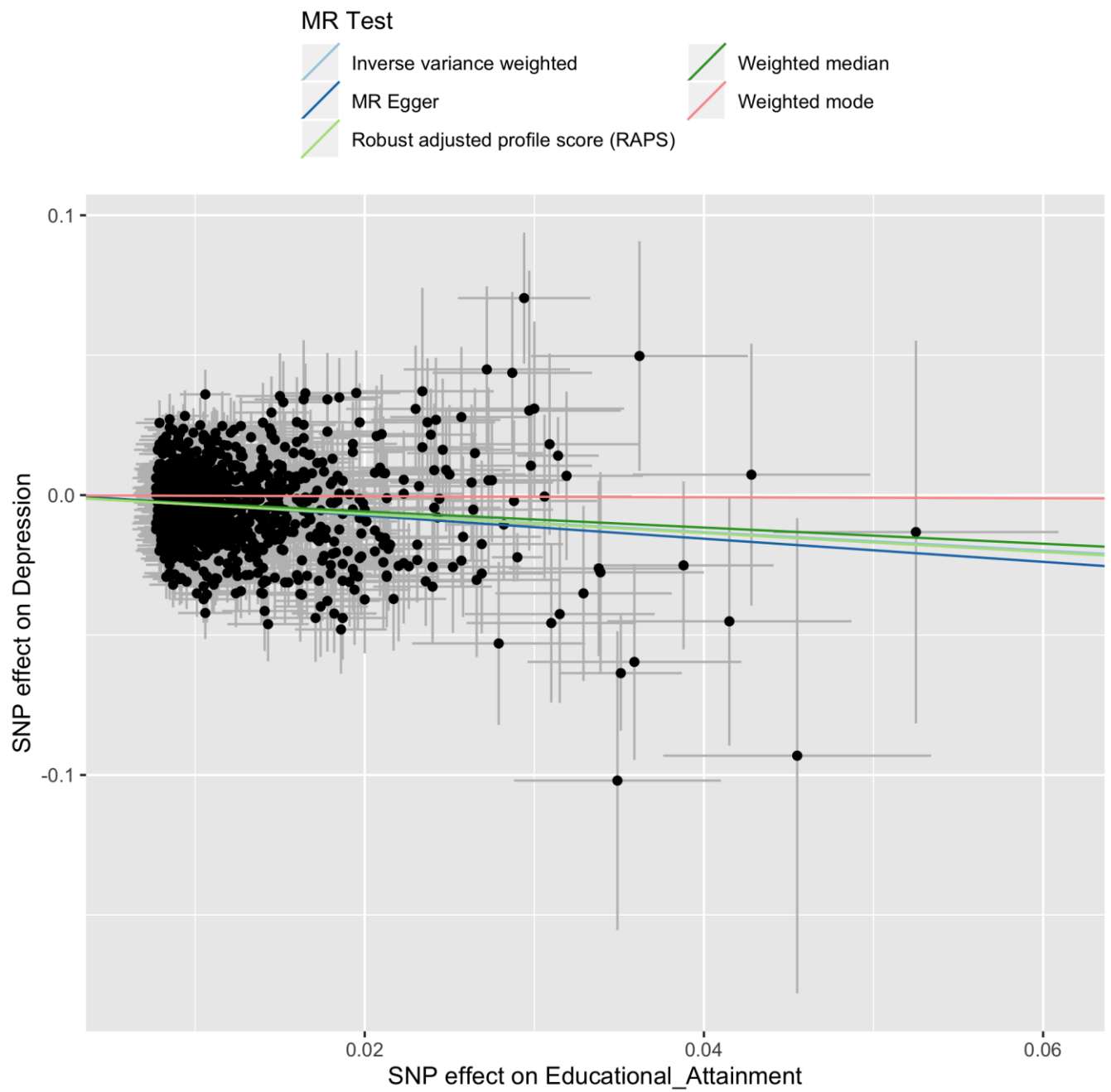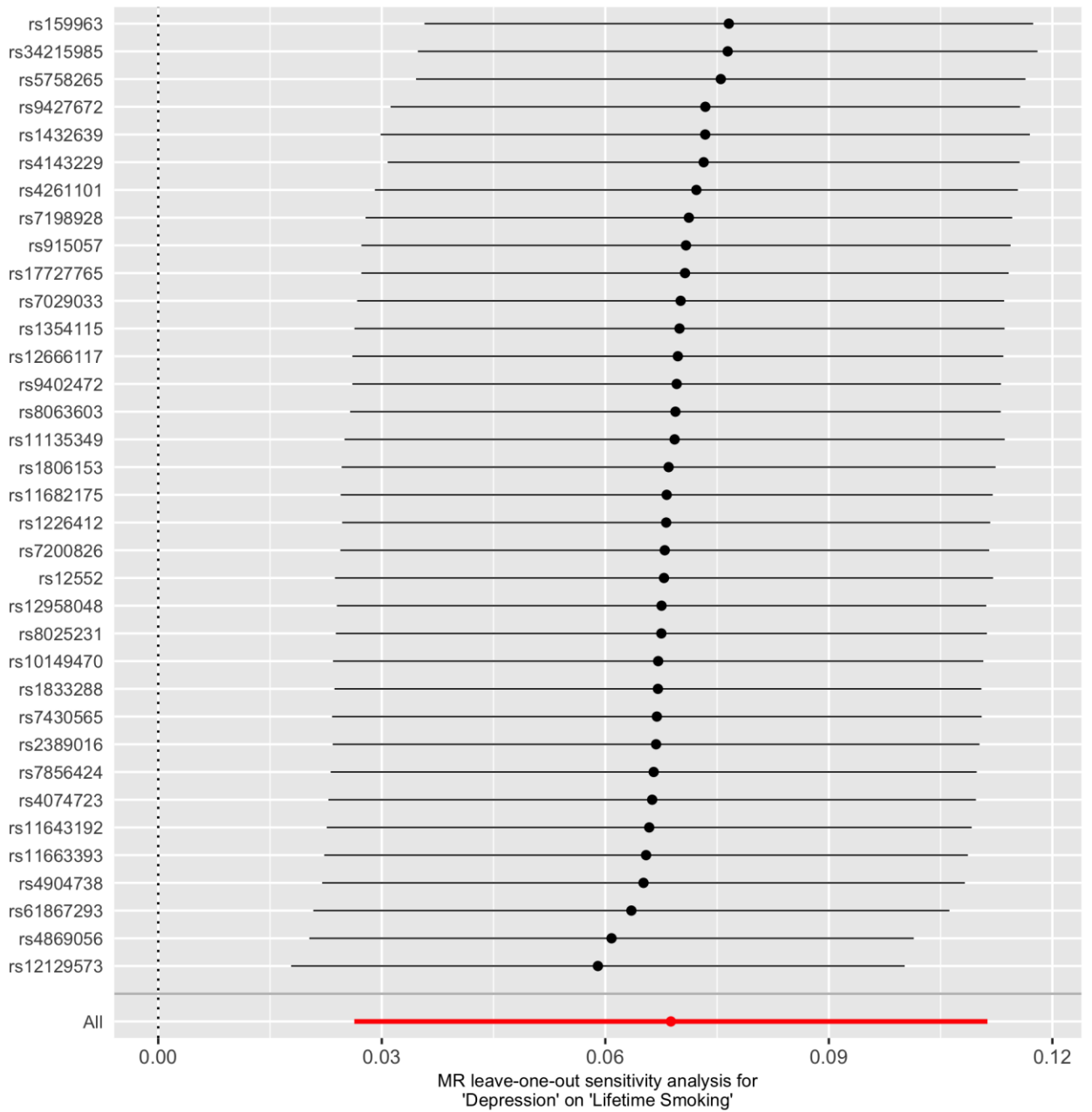**Figure S9. Plot of MR leave-one-out analysis of depression and lifetime smoking.**

MR leave-one-out sensitivity analysis for
'Depression' on 'Lifetime Smoking'

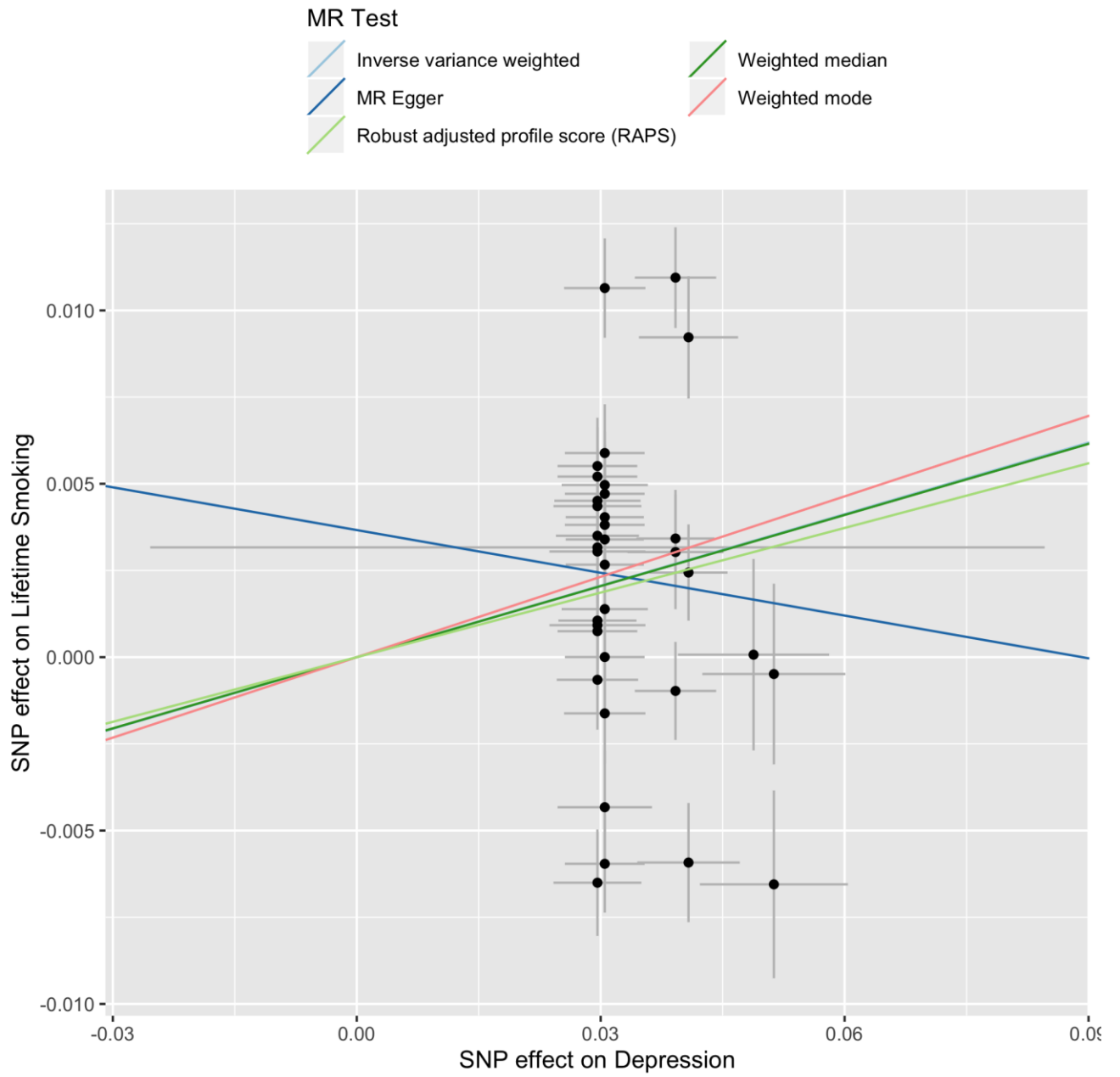**Figure S10. Scatter plot of MR analysis of depression and lifetime smoking.**

Figure S11. Plot of MR leave-one-out analysis of lifetime smoking and cardiovascular disease.

MR leave-one-out sensitivity analysis for
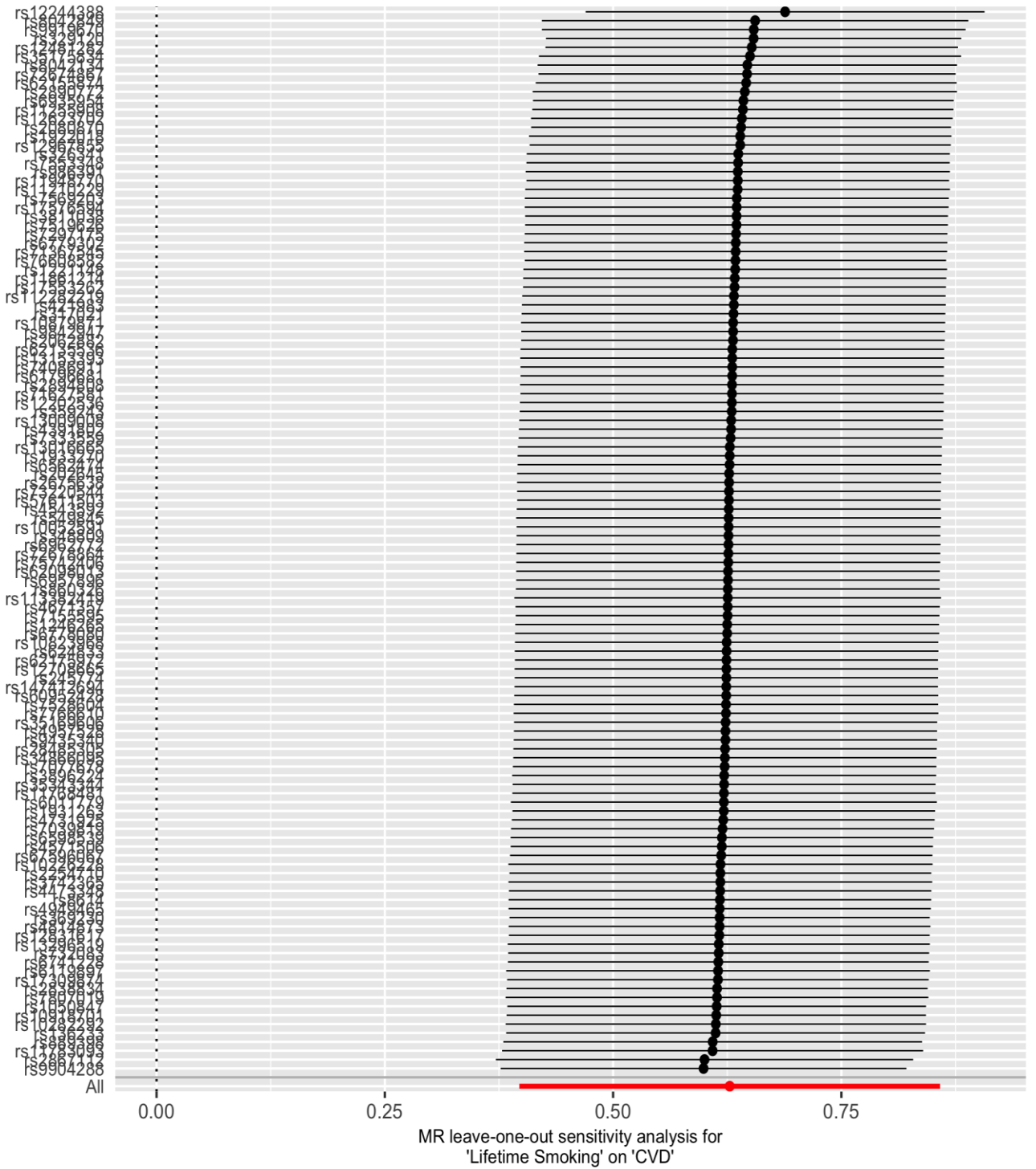'Lifetime Smoking' on 'CVD'

Figure S12. Scatter plot of MR analysis of lifetime smoking and cardiovascular disease.