EXPLORING THE LONG-TERM HEALTH CONSEQUENCES OF ADHD USING A MULTIVARIABLE
MENDELIAN RANDOMIZATION NETWORK APPROACH

A thesis submitted in partial fulfillment of the requirements for the degree of Master of
Sciences in Neurosciences at the Vrije Universiteit

By

Nina S. de Boer, MA

Amsterdam, The Netherlands

Supervisors: Michel G. Nivard, PhD; Sophie van der Sluis, PhD

July 2020

Abstract

Attention-Deficit/Hyperactivity Disorder (ADHD) has been associated with adulthood chronic physical illnesses such as coronary artery disease (CAD) and chronic obstructive pulmonary disease (COPD). However, the potential causes and mediators which underlie this relationship – such as alcohol use, smoking and educational attainment (EA) –, have not yet been studied simultaneously. This study aimed to investigate the relationship between ADHD, CAD, COPD and these potential causes and mediators by construing a direct causal network using genetic information in a multivariable Mendelian randomization (MR) design. After examining how the individual traits were associated to each other by estimating the genetic (partial) correlations, the total and direct causal effects of the traits on each other were examined using univariable and multivariable MR. The analyses demonstrated evidence for the following direct causal pathway: lower EA increases risk of ADHD diagnosis, causing an increased risk of smoking initiation, causing increased COPD risk. No significant causal relationship between ADHD and CAD could be identified. Future studies should address the reliability and validity of the demonstrated relationships.

Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is a condition characterized by inattention, hyperactivity and impulsivity, which interfere with daily functioning (American Psychological Association, 2013). ADHD has a worldwide prevalence of 5.3% in children (Simon et al., 2009) and 2.5% in adults (Polanczyk et al., 2015). Although ADHD typically gets diagnosed in childhood or early adolescence, it is thought to have long-term effects in adulthood going beyond social and psychological disturbances. Indeed, individuals with behavioral disorders (including ADHD) have a higher risk (up to 2.3 times) of developing a variety of physical illnesses later in life (Momen et al., 2020). More specifically, ADHD has been associated with a higher prevalence of coronary artery disease (CAD) (e.g. myocardial infarction) and chronic lung diseases such as chronic obstructive pulmonary disease (COPD) (Smeijns et al., 2013). Adults using ADHD medication are also more likely to receive cardiovascular and respiratory system medications (Zhang et al., 2020). If there is a causal relationship between ADHD and CAD/COPD, this means that ADHD could have adverse health consequences and that interventions should be designed to ameliorate such outcomes in people with ADHD. Before concluding that the consequences are caused by ADHD, however, the association between ADHD and CAD/COPD should be further characterized.

The association between ADHD and CAD/COPD is likely mediated by a variety of factors. First, heavy alcohol use has been associated with ADHD (Brinkman et al., 2015; Charach et al., 2011; Osland et al., 2017) and is considered a risk factor for CAD (Mostofsky et al., 2016; Roerecke et al., 2011; Roerecke and Rehm, 2010). Second, those with ADHD are more likely to be or have been a frequent smoker (Kollins et al., 2005; Milberger et al., 1997; Osland et al., 2017; Pomerleau et al., 1995), and smoking is in turn a risk factor for both CAD (Hackshaw et al.,

2018) and COPD (Pauwels et al., 2001). However, ADHD has also been linked to lower educational attainment (EA) (Polderman et al., 2010), which in turn has been associated with CAD (Kubota et al., 2017; Qureshi et al., 2003), COPD (Gjerdevik et al., 2015), and increased drinking frequency and smoking initiation (Latvala et al., 2014). Lastly, CAD and COPD are also strongly associated, making it difficult to disentangle the contributions of the other factors on these conditions separately (Carter et al., 2019). Hence, when studying the long-term effects of ADHD on these chronic illnesses, the potential mediating roles of alcohol use, smoking and EA should be considered and, ideally, all the potential causes and mediators should be studied simultaneously.

One can conceptualize and represent the relationships between ADHD, EA, smoking, alcohol use, CAD and COPD as a network. There are different types of networks, based on different types of information that can be used to represent the relationships of interest. The most straightforward networks are *correlational* networks that display how traits are associated with each other. Whilst correlational data on these traits is abundantly available, this data cannot demonstrate whether a correlation is attributable to a causal effect or due to confounding. So, *causal* networks that represent the direct, causal effects of all traits on each other have considerably more explanatory power. However, performing randomized controlled trials (RCTs) – the golden standard for establishing causal links between traits– on the entire set of associations between ADHD, EA, smoking, alcohol use, CAD and COPD in order to infer the causal network structure is neither practically feasible nor ethically acceptable.

As an alternative to RCTs, ‘natural’ experiments can be used to draw causal inferences. A prominent example of this is *Mendelian Randomization* (MR), a statistical technique which applies instrumental variable (IV) methods to the field of genetics. These methods are used to test the effect of one trait (the *exposure*) on another trait (the *outcome*). IVs are variables which play a similar role to treatments in an RCT. Hence, valid IVs should meet the following assumptions: they should be 1) strongly associated with the exposure (Relevance assumption), 2) independent of other traits that affect the outcome (Independence assumption) and 3) not affecting the outcome via another path than via the exposure (Exclusion Restriction assumption) (see [Figure 1A](#)) (Angrist et al., 1996; Moffitt, 1996). MR exploits the notion that genetic variants can serve as valid IVs: genetic variants are thought to cause the exposure (and not vice versa), are unconfounded by behavioral, social and physiological factors (Davey Smith et al., 2007), and are subject to little measurement error (Davey Smith and Hemani, 2014). Moreover, Mendel’s second law of random assortment states that the inheritance of genetic variants is *random* (Davey Smith and Ebrahim, 2003). So, whereas RCTs draw causal inferences by examining the differences between a group with and a group without a randomly assigned *treatment*, MR draws causal inferences by examining the differences between a group with and a group without randomly inherited *genetic variants* (Allman et al., 2018).

Nowadays, performing MR research is simplified by the presence of genome-wide association studies (GWASs) that identify genetic variants (usually single nucleotide polymorphisms (SNPs)) statistically associated with a certain trait. MR uses these genome-wide significant SNPs or *hits* as the genetic IV for a trait. If this genetic IV strongly associates with the outcome of interest, there is evidence for a causal effect of the trait (or exposure) on the outcome. Indeed, MR has been used previously to investigate the causal direction of some of the relationships between our traits of interest (e.g. between EA and smoking initiation (Gage et al., 2018) and between EA and alcohol use (Rosoff et al., 2019)).

MR usually assesses the causal relationship between two traits (i.e., *univariable MR*). When drawing causal inferences between more than two traits, MR needs to be carefully extended. First, the assumptions of univariable MR assumptions are often violated: when examining the relations between multiple strongly correlated traits, there is a high possibility that some of the genetic IV SNPs of the exposure affect the outcome via one of the other traits (i.e., *horizontal pleiotropy*), leading to a violation of the Exclusion Restriction assumption (see [Figure 1A](#)). Second, univariable MR cannot assess to what extent the causal relationship between an exposure and outcome is direct or *mediated* by another trait. Mediation and horizontal pleiotropy can be tested or modeled explicitly by making use of *multivariable MR*, which estimates the effect of an exposure on an outcome whilst conditioning on the other exposures (Bowden et al., 2015; Burgess and Thompson, 2015; Sanderson et al., 2019). Indeed, multivariable MR allows all traits in the model to have overlapping genetic IVs and can estimate the unmediated causal effect of an exposure on an outcome ([Figure 1B](#) for the IV assumptions of multivariable MR). So, whereas univariable MR demonstrates the *total* causal effect of an exposure on an outcome, multivariable MR demonstrates the *direct* causal effect.

The main aim of this paper is to study the effect of ADHD on CAD and COPD by construing a direct causal network which includes potential mediators and alternate causes of CAD and COPD: alcohol use, smoking and EA. First, a correlational network will be estimated, after which causal networks with the total and the direct effects will be estimated using univariable and multivariable MR, respectively. Last, these causal networks will be compared to a network estimated using Bayesian inference algorithms.

A graphical overview of prior hypotheses is presented in [Figure 2](#). ADHD is expected to have direct causal effects on alcohol use and smoking, since it is hypothesized that individuals with ADHD smoke and drink more due to their impulse control problems and as a means of self-medication (Wilens et al., 2007). ADHD is also expected to have a direct causal effect on EA, since the symptoms of ADHD may increase the likelihood for problems at school. There is a strong bidirectional association between smoking and alcohol use (Acosta et al., 2008; Bobo and Husten, 2000; Harrison et al., 2008; McKee et al., 2007, 2004; Schorling et al., 1994; Weitzman and Chen, 2005), and it is expected that individuals who smoke are more likely to drink and vice versa. Furthermore, based on previous MR studies, it is hypothesized that lower EA causes increased alcohol use (Rosoff et al., 2019) and higher smoking frequency (Gage et

al., 2018). Although a causal effect of EA on CAD has been demonstrated using MR, it is not hypothesized that this effect is direct, since preliminary evidence demonstrated that it may be mediated by smoking (and other cardiovascular risk factors) (Tillmann et al., 2017). It is hypothesized that smoking directly impacts COPD and CAD risk. Although there is debate on the exact relationship between alcohol use and CAD (e.g. (moderate) alcohol may lead to CAD mortality risk (Dai et al., 2015), but (Millwood et al., 2019)), a causal effect is hypothesized. No direct causal effect between CAD and COPD is hypothesized: although previous MR studies have demonstrated an inverse causal effect of lung function on CAD risk (Au Yeung et al., 2018; Nowak, 2018), these studies did not adjust for the potential confounding effects of smoking.

Since none of the prior hypotheses are based on indisputably strong (unmediated) causal evidence, all (bidirectional) relationships in the model will be investigated. Some relationships are considered implausible due to temporal impossibility (i.e., CAD/ COPD causing ADHD or affecting EA) or extreme unlikeliness (COPD causing smoking). These relationships will still be tested, however, since their presence or absence allows a check of the validity of the model.

Methods

Selection of discovery and replication GWASs

This study used two independent GWASs for every trait: 1) a *discovery* GWAS from which the genome-wide significant SNPs or hits associated with the exposure were obtained (i.e. genetic IV selection), and 2) a *replication* GWAS whose summary statistics were used for the MR and correlational analyses.

This design was used instead of a more straightforward design with only one GWAS per trait because most GWASs had overlapping samples (see Supplementary Table 1 for an overview of the samples which were part of the discovery and replication GWASs and Supplementary Table 2 for the overlap between the discovery GWASs). It has been demonstrated that when the genetic IV and summary statistics are obtained from the same GWAS, sample overlap between exposure and outcome GWASs biases towards finding exposure-outcome associations when the genetic IVs include weak instruments (i.e., SNPs that explain little variation in the exposure) (Burgess et al., 2016). This bias can be prevented by 1) using GWASs with non-overlapping samples, or by 2) using two independent GWASs for both the exposure and the outcome, so that sample overlap becomes a non-issue. Using two independent GWASs for both the exposure and the outcome also decreases the influence of *winner's curse*, i.e., identifying hits whose actual effect sizes are inflated due to random fluctuations (Bowden and Dudbridge, 2009; Haycock et al., 2016). By using an independent GWAS for effect size estimation, the possibility of including SNPs with inflated effect sizes is decreased.

For some traits, one GWAS did not have complete summary statistics publicly available (see Table 1) but only reported the genome-wide hits. Hence, this GWAS was used as discovery GWAS, as in all cases summary statistics for hits (to be used as genetic IVs) could be extracted

from tables in the published data. If both GWASs had summary statistics available, the GWAS with the largest sample size and a population most similar to the other replication GWASs was used as replication GWAS (Zhao et al., 2019). No independent discovery and replication GWAS could be utilized for ADHD, since only one large GWAS on this trait was available (Demontis et al., 2019). Fortunately, no sample overlap was present between this GWAS and the discovery GWASs of the other traits (see [Supplementary Table 1](#)).

[Table 1](#) provides an overview of the operationalizations of each trait, the GWASs used and their sample sizes.

ADHD was operationalized as a clinical diagnosis according to the International Statistical Classification of Diseases and Related Health Problems (ICD10) (World Health Organization, 2004). The ADHD GWAS (Demontis et al., 2019) consisted of twelve cohorts (of which eleven were part of the Psychiatric Genetics Consortium (PGC) and one comprised the iPSYCH-ADHD cohort) with a total of 55,374 individuals including both cases and controls.

Alcohol use was operationalized as the total score of the Alcohol Use Disorders Identification Test – Consumption (AUDIT-C). The AUDIT is a 10-item self-report questionnaire developed by the World Health Organization (WHO) which screens for harmful drinking by assessing alcohol consumption, drinking behaviors and alcohol-related problems (Barbor et al., 2001; Saunders et al., 1993). The Consumption scores refer to the total scoring on the first three items of the AUDIT which measure typical quantity (item 1) and frequency (item 2) of drinking and frequency of heavy or binge drinking (item 3) (maximum score = 12). The discovery GWAS (Kranzler et al., 2019) comprised of 274,424 individuals which were part of the Million Veteran Program. The replication GWAS included individuals which were part of the UK Biobank (analysis ran by TM), comprising 147,267 subjects.

CAD was operationalized as an inclusive CAD diagnosis (e.g. myocardial infarction, acute coronary syndrome, chronic stable angina, or coronary stenosis > 50%). The discovery GWAS (Nikpay et al., 2015) consisted of 48 cohorts and comprised a total of 187,599 individuals including both cases and controls. The replication GWAS included 361,194 individuals (both cases and controls) which were part of the UK Biobank (item ICD10:I25, analysis ran by the Neale Lab).

EA was operationalized as the number of years of schooling that individuals completed. The discovery GWAS was comprised of the previously published EA3 GWAS (Lee et al., 2018) excluding British cohorts and 23&Me, leading to a total inclusion of 53 cohorts with a sample size of 242,952. The replication GWAS was comprised of UK Biobank participants (analysis ran by AA), totaling 451,247 individuals.

COPD was operationalized as the inverse ratio of forced expiratory volume in one second (FEV₁) to forced vital capacity (FVC). This operationalization was chosen instead of a diagnosis because COPD has a low prevalence in population-based cohorts, hampering the possibility of

selecting two independent GWASs with enough COPD cases. Since a low FEV_1/FVC is used for diagnosing COPD (i.e., $\leq 70\%$) (Devine, 2008; Johnson and Theurer, 2014), the inverse ratio was taken as a direct proxy for COPD risk. The discovery GWAS (Soler Artigas et al., 2011) included a total of 40 cohorts and 94,612 individuals and was stratified for smoking status. The replication GWAS was taken from the UK Biobank (item 20154, analysis ran by the Neale Lab), including 117,240 individuals. This sample used a slightly different operationalization, i.e., the ratio of FEV_1 to predicted FEV_1 , a measure which controls for age and posture. Individuals with various lung diseases and “wheezing” were excluded from this GWAS and results were calculated separately for smokers and non-smokers.

Smoking was operationalized as smoking initiation (binary: ever versus never been a regular smoker). Questions in cohorts to assess this include “Have you smoked over 100 cigarettes over the course of your life?”, “Have you ever smoked every day for at least a month?” and “Have you ever smoked regularly?”. As discovery GWAS, the GWAS & Sequencing Consortium of Alcohol and Nicotine use (GSCAN) (Liu et al., 2019) excluding 23&Me and UK Biobank cohorts was used, resulting in a sample comprised of 31 cohorts and a total of 249,171 individuals. The replication GWAS consisted of individuals which were part of the UK Biobank, comprising 301,733 individuals (analysis ran by AA).

Before subsequent analyses, the replication GWASs were pruned to only include SNPs present in all six replication GWASs, in order to avoid the scenario that new proxies for genetic IV SNPs had to be selected for every MR analysis.

Estimation of the (partial) genetic correlation network

To assess how the six traits were genetically correlated, a genetic covariance matrix was estimated. This was done by performing multivariable linkage disequilibrium (LD) score regression using the *genomicSEM* package in *R* (Grotzinger et al., 2019). The summary statistics of all replication GWASs were munged to convert them to the right format for LD score regression. (Replication GWASs were used instead of discovery GWASs since they had larger sample sizes and hence the genetic correlations could be estimated more precisely). Hapmap3 SNPs without the MHC region were used as reference file. Next, multivariable LD score regression was run to estimate a genetic covariance matrix of the six traits. LD scores and weights for populations with a European background were used. Using structural equation modeling (SEM), a model was fitted based on this covariance matrix to obtain the genetic correlations between the individual traits and the standard errors of the genetic correlations. To correct for multiple testing, genetic correlations were considered significant when the false discovery rate (FDR) corrected P value (i.e., Q) < 0.05 .

Next, the partial genetic correlations between the traits were examined – i.e., the genetic associations between pairs of traits whilst conditioning on all other traits in the study. The covariance matrix was scaled to a correlation matrix, after which the off-diagonal elements of the inverse correlation matrix were calculated to obtain the partial correlations. After

calculating the sufficient statistics for an independence test, an order-independent PC-stable algorithm (Colombo and Maathuis, 2014) was used to estimate the significant partial correlations ($\alpha = 0.05$).

Selection of the genetic IVs and preprocessing of summary statistics for MR

To assess how the six traits were causally related, MR was performed. To this end, genetic IVs were selected for every trait. The following steps were performed for all six exposures.

First, independent hits were obtained from the discovery GWAS. If a discovery GWAS did not report summary statistics (see [Table 1](#)), hits were extracted from the available publication. If a publication reported a genome-wide significant insertion/deletion and an associated SNP ($R^2 \geq 0.8$), the associated SNP was selected. If summary statistics were available, genome-wide significant SNPs ($P < 5 \times 10^{-8}$) were selected from the GWAS using the *MR-Base* package in *R* (Hemani et al., 2018). Clumping was performed to identify independent hits: clumping identifies independent signals by considering the LD between SNPs against a reference dataset of similar ancestry and selects the SNP which is most strongly associated with the exposure (clumping r^2 cutoff = 0.001) (Hemani et al., 2018).

Second, the independent hits were identified in the replication GWAS. If a SNP identified in the discovery GWAS was absent in the replication GWAS, or if the SNP was palindromic (i.e., had the same allele pair on the forward as on the reverse strand), a proxy was identified using the LDproxy function from the *LDlinkR* package in *R* (Machiela and Chanock, 2015). Proxies were only included in the genetic IV when strongly associated to the original hit ($R^2 \geq 0.8$).

Third, the effect size (β), standard error (S.E.) and P value of the genetic IV SNPs were extracted from the replication GWAS. If a replication GWAS reported odds ratios, these were converted to beta coefficients by calculating the logarithm of the odds ratio. The final genetic IVs only included SNPs which associated with the exposure in question with $P < 0.05$ in the replication GWAS (Wang et al., 2020). (Note that it was required that the genetic IVs contained at least as many SNPs as there were exposures in the model (in this case, five), since this is a prerequisite for multivariable MR analyses (Sanderson et al., 2019).)

Finally, the SNP-outcome associations were obtained by extracting the β , S.E. and P value of the final genetic IV SNPs from the replication GWASs of the outcome traits. SNP alleles were harmonized to ensure that the effect allele and the other allele were similarly specified in every replication GWAS (Hartwig et al., 2016).

Estimation of the causal network using univariable and multivariable MR

All MR analyses were performed using the *MendelianRandomization* package in *R* (Yavorska and Burgess, 2017), unless specified otherwise.

Total causal network (univariable MR)

To assess how the six traits were causally related without controlling for possible mediating effects (i.e., the total causal effects), univariable MR was performed. The size of the total causal effect was estimated using the *inverse-variance weighted (IVW) method* with multiplicative random effects. This method calculates the Wald ratio (SNP-outcome association / SNP-exposure association (Boef et al., 2015; Lawlor et al., 2008)) for each genetic IV SNP and pools the results of all SNPs using an IVW meta-analysis approach (Johnson, 2012; Thompson et al., 2016). IVW estimates were considered significant when $Q < 0.05$.

To assess the strength of the statistical evidence for the IVW estimates, three other causal estimation methods were used. First, *MR-Egger regression* (Bowden et al., 2015) relaxes the Exclusion Restriction assumption (see [Introduction](#)), returning an unbiased causal estimate even when this assumption is violated for all genetic IV SNPs. It calculates the Wald ratio for each genetic IV SNP and pools these ratios using an adapted Egger regression which allows a non-zero intercept. The regression slope coefficient is interpreted as the causal effect of the exposure on the outcome. MR-Egger regression does assume that the influence of a SNP on a confounding exposure is not correlated with the size of the SNP-exposure effect (InSIDE assumption) (Bowden et al., 2015). Second, the *weighted median-based method* provides a consistent estimate of the causal effect when half of the genetic IV SNPs is invalid (i.e., violates one or more of the IV assumptions). This method first uses the IVW method to obtain the Wald ratio and weights for each genetic IV SNP, after which the median of the causal estimates is calculated (Bowden et al., 2016; Kang et al., 2016). Last, the *weighted mode-based method* also gives consistent results when part of the genetic IV SNPs is invalid. It obtains the mode of the causal estimates of every SNP by clustering them into groups based on similarity of causal effects and returns the causal effect estimate based on the cluster with the largest number of SNPs (Hartwig et al., 2017). The causal estimates of these three techniques were considered significant when $P < 0.05$. Since IVW is the most straightforward univariable MR method, causal effects not significant according to IVW will not be discussed.

Additional analyses were used to assess the robustness of the total causal effect estimates. Instrument strength was assessed using Cochran's Q statistic (Bowden et al., 2019; Cochran, 1954; Del Greco M et al., 2015) of both the IVW method and MR-Egger regression (i.e., Cochran's Q_{IVW} and Q_{EGGER}). A significant Cochran's Q statistic indicates high heterogeneity between the genetic IV SNPs, which means that one or more SNPs may violate IV assumptions, for example due to horizontal pleiotropic effects. The presence of directional pleiotropy (i.e., pleiotropic effects of SNPs on the outcome all or mostly acting in one direction to increase or decrease the outcome) was assessed using the intercept of MR-Egger regression (Bowden et al., 2015). Significance for the robustness measures was set at $P < 0.05$.

Direct causal network (multivariable MR)

To assess how the six traits were causally related whilst controlling for potential mediating effects (i.e., the direct causal effects), multivariable MR was performed. The size of the direct causal effect was estimated using the *multivariable IVW method*, which regresses the SNP-outcome associations on all SNP-exposure associations using a multivariable weighted regression model (Burgess and Thompson, 2015). IVW estimates were considered significant when $Q < 0.05$. To assess the strength of the statistical evidence for the multivariable IVW estimates, *multivariable MR-Egger regression* (Rees et al., 2017) was used (i.e., a version of MR-Egger which is able to accommodate for the presence of multiple variables). Multivariable MR-Egger estimates were considered significant when $P < 0.05$. Since multivariable IVW is the most straightforward multivariable MR method, causal effects not significant according to multivariable IVW will not be discussed.

Additional analyses were used to assess the robustness of the direct causal effect estimates. Instrument strength (i.e., how strong do genetic IV SNPs predict each trait in the model) was assessed using the *conditional F statistic* (Sanderson et al., 2020). If $F < 10$, there is evidence for weak instrument bias. The presence of horizontal pleiotropy was evaluated using a modified version of Cochran's Q statistic which has an appropriate Type 1 error rate when weak instruments are present: Cochran's Q_A (Sanderson et al., 2020). A significant Cochran's Q_A indicates potential horizontal pleiotropy in the form of excessive heterogeneity. To estimate the F and Cochran's Q_A statistic, the *MVMR* package in *R* was used (Sanderson et al., 2019). The presence of directional pleiotropy was assessed using the intercept of multivariable MR-Egger regression (Rees et al., 2017). Significance for the robustness measures was set at $P < 0.05$.

Estimation of the directed acyclic graph (DAG)

To triangulate the direct causal network with another causal inference technique, a DAG was estimated using Bayesian network inference algorithms. This DAG described the causal relationships between the traits of interest based on correlational data and was used as a comparison to the direct causal network to examine its reliability. Estimation was based on the genetic covariance matrix obtained using LD score regression, and the undirected significant partial correlations served as the DAG's skeleton (see Estimation of the partial (genetic) correlation network). The final DAG (i.e., with directed edges) was estimated using a conditional independence test ($\alpha = 0.05$).

All networks were plotted using the *qgraph* package in *R* (Epskamp et al., 2012).

Results

Pruning the replication GWASs resulted in a total of 6.224.695 SNPs present in every replication GWAS (77% of the original number of SNPs in the replication GWAS of ADHD, 72% of alcohol use, 45% of CAD, 59% of EA, 45% of COPD and 59% of smoking).

Correlations

Of the 15 correlation estimations, 12 were significant (75%). The negative genetic association between ADHD and EA was the strongest in the network ($r_g = -0.489$, S.E. = 0.027, $Q = 4.76 \times 10^{-71}$), and both ADHD and EA were significantly genetically associated with all other traits. ADHD was positively genetically associated with smoking initiation ($r_g = 0.237$, S.E. = 0.037, $Q = 3.46 \times 10^{-9}$) and CAD ($r_g = 0.193$, S.E. = 0.044, $Q = 1.08 \times 10^{-4}$) and COPD ($r_g = 0.169$, S.E. = 0.033, $Q = 4.86 \times 10^{-6}$) and alcohol use ($r_g = -0.107$, S.E. = 0.039, $Q = 0.038$), EA was negatively genetically associated with these respective traits (smoking initiation: $r_g = -0.117$, S.E. = 0.026, $Q = 6.81 \times 10^{-5}$; CAD: $r_g = -0.225$, S.E. = 0.029, $Q = 6.07 \times 10^{-13}$; COPD: $r_g = -0.188$, S.E. = 0.022, $Q = 2.31 \times 10^{-16}$; alcohol use: $r_g = 0.195$, S.E. = 0.029, $Q = 7.85 \times 10^{-10}$). Moreover, smoking initiation and alcohol use were strongly positively genetically correlated ($r_g = 0.358$, S.E. = 0.040, $Q = 1.38 \times 10^{-17}$), as were CAD and COPD ($r_g = 0.161$, S.E. = 0.037, $Q = 1.30 \times 10^{-4}$).

Of the 15 partial genetic correlation estimations, eight were significant (53%). The strongest genetic correlations – between ADHD and EA and between smoking initiation and alcohol use – remained present when conditioning on the other traits in the model (respectively $pr_g = -0.412$ and $pr_g = 0.408$, $P < 0.05$). Again, ADHD and EA were the most connected traits in this network: ADHD was positively genetically associated to smoking initiation ($pr_g = 0.234$, $P < 0.05$) and COPD ($pr_g = 0.063$, $P < 0.05$), and EA was negatively genetically correlated to CAD ($pr_g = -0.151$, $P < 0.05$) and COPD ($pr_g = -0.093$, $P < 0.05$). Last, a positive partial genetic correlation between CAD and COPD was demonstrated ($pr_g = 0.122$, $P < 0.05$).

The (partial) genetic correlations between the traits are presented in [Table 2](#). The significant genetic (partial) correlation networks are presented in [Figure 3A](#) and [3B](#).

Genetic IV selection

The independent hits obtained from the discovery GWASs and the final genetic IVs obtained from the replication GWASs are presented in [Supplementary Table 3](#). The summary statistics for ADHD, alcohol use, CAD, EA, COPD and smoking in the replication GWASs are presented in [Supplementary Table 4](#). For more information on genetic IV selection, see [Supplementary Information](#).

Univariable MR: total causal effects

Of the 30 IVW estimates, six were significant when correcting for multiple testing (20%). [Figure 3C](#) presents the significant total (IVW) causal effects network and [Figure 3E](#) presents the

number of estimation methods supporting these findings. The causal estimates of the four estimation methods (i.e., IVW, MR-Egger, weighted median and weighted mode) are presented in [Table 3](#), and the robustness measures assessing heterogeneity and directional pleiotropy are presented in [Supplementary Table 5](#).

ADHD

Evidence for a negative total causal effect of ADHD on EA was demonstrated ($\beta_{IVW} = -0.428$, S.E. = 0.119, $Q = 0.002$), although this effect could only be corroborated by the weighted median method. No evidence for directional pleiotropy was demonstrated, but there was evidence for heterogeneity (Cochran's $Q_{IVW} = 79.016$, $P = 7.83 \times 10^{-13}$; Cochran's $Q_{EGGER} = 72.295$, $P = 5.39 \times 10^{-12}$). Evidence for a causal effect of ADHD on smoking initiation was also demonstrated ($\beta_{IVW} = 0.018$, S.E. = 0.005, $Q = 0.002$), a finding supported by the weighted median and mode method. No evidence for directional pleiotropy or heterogeneity was demonstrated.

Smoking initiation

Evidence for a total causal effect of smoking initiation on ADHD status was shown ($\beta_{IVW} = 11.982$, S.E. = 3.679, $Q = 0.006$), a finding supported by the weighted median and mode method. There was evidence for heterogeneity (Cochran's $Q_{IVW} = 26.406$, $P = 7.44 \times 10^{-5}$), but no evidence for directional pleiotropy. Evidence for a total causal effect of smoking initiation on COPD was also demonstrated ($\beta_{IVW} = 3.697$, S.E. = 0.752, $Q = 1.31 \times 10^{-5}$) and corroborated by both the weighted median and mode method. No evidence for directional pleiotropy was demonstrated, but there was evidence for heterogeneity (Cochran's $Q_{IVW} = 11.388$, $P = 0.044$).

Alcohol use

There was evidence of a total causal effect of alcohol use on COPD risk ($\beta_{IVW} = 0.922$, S.E. = 0.238, $Q = 1.06 \times 10^{-3}$). This finding was supported by all other analysis methods, and there was no evidence for directional pleiotropy or heterogeneity.

EA

There was evidence that EA exerts a negative total causal effect on ADHD status ($\beta_{IVW} = -0.221$, S.E. = 0.028, $Q = 2.22 \times 10^{-13}$), a finding supported by all analysis methods. No evidence for directional pleiotropy or heterogeneity was demonstrated.

CAD

Evidence of a total causal effect of CAD on smoking initiation was observed, but this effect was not significant when correcting for multiple testing ($\beta_{IVW} = 0.182$, S.E. = 0.072, $P = 0.012$, $Q = 0.051$). This finding was not supported by other analysis methods and there was evidence

for heterogeneity (Cochran's $Q_{IVW} = 68.649$, $P = 5.81 \times 10^{-4}$ and Cochran's $Q_{EGGER} = 60.943$, $P = 0.003$) and directional pleiotropy (intercept estimate = 9.64×10^{-4} , $P = 0.038$).

COPD

There was evidence for a total causal effect of COPD on CAD status, but only when not correcting for multiple comparisons ($\beta_{IVW} = 0.031$, S.E. = 0.014, $P = 0.030$, $Q = 0.112$). This finding was not supported by other analysis methods and there was evidence for heterogeneity (Cochran's $Q_{IVW} = 28.470$, $P = 3.93 \times 10^{-4}$, Cochran's $Q_{EGGER} = 28.319$, $P = 1.92 \times 10^{-4}$).

Multivariable MR: direct causal effects

Of the 30 multivariable IVW estimates, two were significant when correcting for multiple testing (7%). [Figure 3D](#) presents the significant direct (multivariable IVW) causal effects network and [Figure 3F](#) presents the number of estimation methods supporting these findings. The causal estimates of the two estimation methods (i.e., multivariable IVW and multivariable MR-Egger) are presented in [Table 4](#), and the robustness measures assessing heterogeneity and directional pleiotropy are presented in [Supplementary Table 6](#).

ADHD

There was evidence for a direct causal effect of ADHD on smoking initiation, but this effect was not significant when correcting for multiple comparisons ($\beta_{MVIW} = 0.021$, S.E. = 0.010, $P = 0.042$, $Q = 0.318$). It was supported by multivariable MR-Egger, however. Although there was no evidence for horizontal pleiotropy, there was evidence for weak instruments (all $F < 0.97$) and directed pleiotropy (intercept estimate = -0.004, $P = 0.025$). The previously identified total causal effect of ADHD on EA was non-significant.

Smoking initiation

Evidence for a direct causal effect of smoking initiation on COPD risk was demonstrated, but this finding was non-significant when controlling for multiple comparisons ($\beta_{MVIW} = 3.889$, S.E. = 1.727, $P = 0.024$, $Q = 0.243$). Multivariable MR-Egger could not be performed due to the low number of genetic IV SNPs for smoking initiation. Evidence for weak instruments (all $F < 1.5$) and horizontal pleiotropy (Cochran's $Q_A = 1.001$, $P = 0.000$) was demonstrated. The previously identified total causal effect of smoking initiation on ADHD was non-significant.

Alcohol use

There was evidence for a direct causal effect of alcohol use on COPD ($\beta_{MVIW} = 1.114$, S.E. = 0.249, $Q = 1.16 \times 10^{-4}$), a finding corroborated by multivariable MR-Egger. There was evidence for weak instruments (all $F < 1.3$), but not for pleiotropy.

EA

Evidence for a direct causal effect of EA on ADHD was presented ($\beta_{\text{MIVW}} = -0.209$, S.E. = 0.032, $Q = 2.01 \times 10^{-9}$), although this effect was not corroborated by multivariable MR-Egger. There was evidence for weak instruments (all $F < 4.8$), but no evidence for pleiotropy.

CAD

No evidence for significant direct causal effects of CAD on other traits were demonstrated (including the previously identified total causal effect on smoking initiation).

COPD

No evidence for significant direct causal effects of COPD on other traits were demonstrated (including the previously identified total causal effect on CAD).

Directed Acyclic Graph

Figure 3G presents the estimated DAG. A causal effect of ADHD on smoking initiation was demonstrated. Although no causal effects of smoking initiation did not have causal effects according to the DAG, a causal effect of alcohol use on smoking initiation and EA was presented. It showed a causal effect of were presented, it was shown that alcohol use causes both smoking initiation and EA and that EA was causally related to ADHD. Causal effects of CAD and COPD on EA were demonstrated, as well as a causal effect of COPD on ADHD. Last, the direction of the edge between CAD and COPD could not be estimated.

Summary of main findings

Table 5 shows a summary of the significant correlational and causal relations and how these relate to the prior hypotheses. The negative causal effect of EA on ADHD was the only relation which was significant according to all main estimation methods.

Discussion

ADHD has been associated with adulthood chronic physical illnesses such as CAD and COPD; however, the potential causes and mediators which underlie this relationship have not yet been studied simultaneously. To this end, this study investigated the relationship between ADHD and CAD/COPD and the potential (mediating) roles of alcohol use, smoking and EA using genetic information in an MR design to establish which relationships are causal. It was first examined how the individual traits were associated to each other by estimating the genetic correlations, and how they were associated whilst conditioning on the other traits by estimating the genetic partial correlations. Subsequently, the total and direct causal effects of the traits on each other were examined using univariable and multivariable MR, respectively. The direct causal estimates were in turn triangulated with a DAG estimated using Bayesian causal inference.

The genetic correlation network shows that ADHD genetically correlates to all traits under investigation: having ADHD is genetically associated with lower EA, increased substance use and a higher risk of adulthood physical illness. Whilst the genetic correlations corroborate phenotypical evidence, the correlational network does not show how individual traits are *directly* associated. The partial correlation network demonstrates that the genetic relation between ADHD and CAD is indirect: EA and COPD are involved in the association. It also demonstrates that the relation between ADHD and COPD may be direct, but that EA could also be involved in the association.

In line with the hypotheses, the partial genetic correlation network shows a strong negative genetic association between ADHD and EA. However, it also shows that alcohol use and smoking initiation are genetically associated to ADHD, but not implicated in the relationship between ADHD and adulthood physical illness, which is contrary to expectations. Although a previous MR study demonstrated that an ADHD diagnosis does not lead to increased alcohol consumption, it was demonstrated that it leads to a higher chance of smoking initiation (Treur et al., 2019). Moreover, the demonstrated direct associations between adulthood physical illnesses and both ADHD and EA are unexpected.

Contrary to the partial genetic correlation network, the total causal effects network suggests that smoking initiation could mediate the relation between ADHD and adulthood physical illness: ADHD causes an increased smoking initiation risk, which in turn increases one's COPD risk, which in turn increases one's CAD risk. There is strong evidence for the first two causal links, since they are supported by IVW as well as two more univariable MR analysis methods. However, the causal effect of COPD on CAD is small, does not survive correction for multiple testing and is only supported by the IVW method. Surprisingly, the total causal network highlights the possibility that the causal effects of ADHD on adulthood physical illness are in fact primarily due to lower EA, which increases one's risk of having ADHD.

The direct causal effects network demonstrates that the relation between ADHD and COPD corroborates the total causal effects: lower EA causes an increased risk of ADHD diagnosis, which increases smoking initiation risk, which increases one's risk for COPD. Notably, only the effect of EA on ADHD survives correction for multiple testing. When examining the direct causal relation between ADHD and CAD, no connection is visible in the network. In fact, none of the traits in the direct causal network have a direct causal effect on CAD.

Although the direct causal effects network corroborates the hypotheses that ADHD causes smoking initiation, and that smoking initiation causes COPD, these effects do not survive correction for multiple testing. Moreover, the other hypotheses are unexpectedly not supported by the model. First, the direct causal network shows that EA causes ADHD. Although the association between ADHD and EA has been frequently reported, it is unlikely that the latter causes the former: ADHD is generally diagnosed in childhood or adolescence, i.e., before finishing education. Although it is possible that getting tested for ADHD is biased by educational performance (e.g., children with lower educational performance are more frequently tested on ADHD to explain their 'poor performance'), this finding is not intuitive. Second, the direct causal effects network does not demonstrate that alcohol use mediates the relationship between ADHD and adulthood physical illness, but instead demonstrates a strong direct causal effect of alcohol use on COPD risk. Although there is correlational evidence that high alcohol consumption is associated with lung function and COPD risk (Tabak et al., 2001), some studies report no effect (Greene et al., 2008), and some studies even indicate a positive association between light to moderate drinking and lung function (Siu et al., 2010). To date, no studies have investigated the causal effect of alcohol use on COPD. Third, it is unlikely that none of the traits in the model exert a causal effect on CAD.

These unexpected findings may indicate that the direct causal network has limited reliability and/or validity. To assess its reliability, the direct causal network can be compared to other lines of evidence. First, comparing the direct causal network to the partial genetic correlation network demonstrates that causal effects are present where no partial correlations were observed (i.e., the causal effects of both smoking initiation and alcohol use on COPD). This finding is counterintuitive: if two traits are causally connected, they should be necessarily correlated to each other. However, the absence of a (partial) correlation does not imply that this correlation does not exist; rather, it means that this connection was not *significant*. The discrepancy can also be explained by the differences in the methods used: MR uses only a few SNPs, whereas genetic correlation methods use the whole GWAS. Second, comparing the direct causal network to the DAG estimated using Bayesian network inference shows striking differences: instead of finding causal effects of ADHD on CAD/COPD, the DAG demonstrates that these adulthood physical illnesses cause ADHD, either directly or indirectly via EA. Since these causal pathways are temporally impossible (or highly implausible), it can be argued that this DAG is an unreliable comparison. Indeed, DAGs are not able to account for bidirectional and cyclic relationships, and empirical studies have demonstrated that Bayesian network estimations that do not include SNPs as instruments do not perform better than chance

(Howey et al., 2020). So, the reliability of the direct causal network cannot be tested in a satisfactory manner by comparing it to the partial correlation network or to the estimated DAG.

The validity of the directed causal network can be assessed by systematically evaluating whether the genetic IV assumption for multivariate MR have been met – if these assumptions have been violated, the results cannot be fully trusted. First, the Relevance assumption states that the genetic IV should be strongly associated to the exposure(s) in the model. However, the robustness analyses demonstrated that this assumption is violated: all multivariable MR comparisons show evidence for the presence of weak instruments (see [Supplementary Table 6](#)). This could be due to relying on instrument selection in the discovery GWAS according to genome-wide significance and then relaxing the instrument inclusion criterion to $P < 0.05$ in the replication GWAS. Indeed, it has been argued that a genome-wide significance threshold should be used to avoid weak instrument bias (Burgess and Thompson, 2011). The $P < 0.05$ threshold was chosen to prevent the reintroduction of winner's curse in the replication GWAS and to ensure that the number of genetic IV SNPs was larger than the number of exposures included in the MR model (a prerequisite of multivariable MR). Indeed, if a more stringent threshold would have been used, the genetic IVs of both smoking initiation and COPD would not have had enough SNPs to perform multivariable MR (see [Supplementary Table 7](#)).

But are weak instruments necessarily problematic? It has been demonstrated that weak instruments can be informative and that genetic IVs that combine weak and strong SNPs do not harm MR and sometimes even increase their accuracy and stability by increasing the power of the genetic IV (Wang et al., 2020). Considering the multivariable MR IV assumption, it has been argued that it is inevitable that genetic IV SNPs do not strongly associate with all traits in a model, and that including weak instruments makes multivariable MR more accurate (Wang et al., 2020). Whether this claim also holds for genetic IVs with few SNPs, is up for discussion. Nonetheless, the presence of weak instruments would not introduce spurious exposure-outcome association, but rather increase the potential of false negatives (Sanderson et al., 2020). So, although the Relevance assumption was likely violated, it is unlikely that it would have explained the unexpected effects of EA on ADHD and of alcohol use on COPD.

Second, the Independence assumption states that there cannot be anything which causes both the genetic IV and the outcome, and its violation could lead to spurious exposure-outcome associations. Although this cannot be investigated directly, it is possible that this assumption was violated for EA due to the presence of *dynastic effects*, i.e., the expression of parental genotype in the parental phenotype that directly affects the offspring's phenotype (Brumpton et al., 2019). Indeed, higher educated parents may support their children's education, causing them to obtain higher EA. Moreover, it has also been demonstrated that having a parent with lower EA is associated with a higher risk of ADHD (Torvik et al., 2020). So, a violation of this assumption via dynastic effects may have contributed to the demonstrated unexpected causal effect of EA on ADHD.

Third, the Exclusion Restriction assumption states that the association of the exposure to the outcome should not be mediated by other traits. Many of the univariable MR estimates show evidence of pleiotropy, validating the use of multivariable MR methods that relax this assumption. Indeed, many demonstrated total causal effects become non-significant in the direct causal network, indicating that they are likely due to horizontal pleiotropic (direct) effects of the exposures on the outcomes via other traits in the model. However, while multivariable MR explicitly models horizontal pleiotropy, evidence for pleiotropy is still observed in the direct causal network. This means that some of the demonstrated relationships are likely to be mediated by confounding factors which were not included in the model. However, the direct causal relationship between EA and ADHD and between alcohol use and COPD did not demonstrate evidence of pleiotropy.

In sum, it could be possible that the unexpected results are due to potential violations of the IV assumptions, especially concerning the causal effect of EA on ADHD. Future studies should examine whether the current results hold when potential dynastic effects are circumvented (for example, by using a within-family MR design (Brumpton et al., 2019)) and when currently unmeasured traits are included in the model (to account for pleiotropy). Moreover, to serve as a more reliable method for causal inference, future studies should compare the multivariable MR findings to Bayesian network estimations that include SNPs as instruments, since these models were shown to outperform chance (and in some instances even MR) (Howey et al., 2020).

Finally, there are other limitations to this study that can be put forward. First, the claims concerning COPD should be interpreted with caution. Although the inverse FEV₁/FVC can be considered a good proxy for COPD, future studies should elucidate whether the same results will be obtained when COPD diagnosis is used in the model. Since ADHD is also associated with other chronic lung diseases (e.g. asthma and chronic bronchitis) (Semeijn et al., 2013), future studies should investigate whether the same relations hold when these diseases are included in the model. The operationalizations of the COPD discovery and replication GWAS were also slightly different, which may have led to weaker effect size estimates. Second, it should be noted that for every trait, one specific operationalization was chosen. This is especially salient for smoking and alcohol use, since both traits can be operationalized in a variety of ways. For instance, the number of cigarettes smoked per day and the ability to quit smoking may be differentially related to ADHD and adulthood physical illnesses than smoking initiation. Similarly, propensity for alcohol use disorder could be differentially related to these conditions than alcohol consumption (as previous MR studies have indicated that ADHD increases one's risk of alcohol dependency (Treur et al., 2019)). Future studies should assess whether the currently identified causal pathways hold when different operationalizations for these traits are used. Last, it should be noted that although MR can help evaluating the health consequences of ADHD, its findings should be critically evaluated and triangulated with other types of evidence.

Conclusion

This study investigated the relationship between ADHD and CAD/ COPD and the potential causal and mediating roles of alcohol use, smoking and EA. It demonstrated evidence for the following direct causal chain of effects: lower EA increases one's risk for ADHD diagnosis, which increases one's risk of smoking initiation, which increases one's risk of developing COPD. No (in)direct causal relationship between ADHD and CAD could be identified given the variables selected for this study. To evaluate the consistency of the direct causal network, it was compared to networks estimated using Bayesian network inference and partial genetic correlations. However, triangulation between these methods remained somewhat unsatisfactory. The (multivariable) MR findings are only valid if the assumptions of the model hold, and it is possible that in some instances, the IV assumptions were violated. Future studies are required to replicate and validate the current findings.

As has been argued previously: "Drawing valid causal inferences on the basis of observational data is not a mechanistic procedure but rather always depends on assumptions that require domain knowledge and that can be more or less plausible." (Rohrer, 2018, p. 1). Using (multivariable) MR in psychiatric epidemiology holds the promise of laying bare causal chains that explain the adverse health outcomes related to disorders such as ADHD. However, causal inference based on natural experiments requires constant (re)evaluation and critical assessment of the results considering the assumptions and other lines of evidence.

Acknowledgements

First, I would like to thank both Michel Nivard and Sophie van der Sluis for the helpful feedback on my manuscript, and Michel specifically for the daily supervision and guidance. I would also like to thank Abdel Abdellaoui (Amsterdam UMC, AMC) for providing the discovery and replication GWAS for EA and the replication GWAS for smoking initiation. Relatedly, I would like to thank Scott Vrieze (University of Minneapolis) for providing the discovery GWAS for smoking initiation, and Travis Mallard and Paige Harden (UTexas) for providing the replication GWAS for alcohol use. Lastly, I would like to thank Perline Demange for her help and for proofreading the manuscript.

Figure 1: The main instrumental variable (IV) assumptions of univariable (A) and multivariable (B) Mendelian Randomization (MR). Red lines indicate relationships that should not be present. The Relevance assumption (1): the IV should be strongly associated with the exposure (in multivariable MR: exposures). The Independence assumption (2): the IV should be independent of other traits that affect the outcome. The Exclusion Restriction assumption (3): the association between the IV and the outcome should be entirely mediated by the exposure (in multivariable MR: exposures). Note that this figure only presents two exposures for multivariable MR, but that more than two can be studied.

Figure 2: The hypothesized causal network. ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, SM = Smoking.

Figure 3: Estimated networks of the six traits. ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, SM = Smoking.

A & B. Each edge corresponds to a significant association (false discovery rate (FDR) corrected $Q < 0.05$) between two traits. The thickness of an edge represents strength, the color represents direction (green = positive, red = negative)

C & D. Each arrow corresponds to a total (C) or direct (D) causal relationship (Z-scored effect size) between two traits, assessed using the inverse-variance weighted method (IVW, C) and multivariable inverse-variance weighted method (MVIVW, D). Solid lines represent effects which were significant using $Q < 0.05$, dashed lines represent effects which were only significant using an uncorrected $P < 0.05$. The thickness of an arrow represents strength, the color represents direction (green = positive, red = negative).

E & F. Each arrow corresponds to a significant total (E) or direct (F) causal relationship ($P < 0.05$) between two traits, assessed using the inverse-variance weighted method (IVW, E) and multivariable inverse-variance weighted method (IVW, F). The thickness and color of an arrow represent the number of estimation methods which also demonstrate a significant causal effect: the thicker/ more blue, the more methods corroborating this finding. The techniques included are IVW, MR-Egger regression, mode-based estimation and median-based estimation for (E), and MVIVW and multivariable MR-Egger regression for (F).

G. Directed Acyclic Graph (DAG) of the causal effects based on Bayesian network inference. Arrows indicate estimated causal relationships ($\alpha = 0.05$) between two traits.

Table 1: Characteristics of genome-wide association studies (GWASs) and operationalizations used for every trait. ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, AUDIT-C = Alcohol Use Disorders Identification Test – Consumption, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, GSCAN = GWAS & Sequencing Consortium of Alcohol and Nicotine use, PGC = Psychiatric Genetics Consortium, SM = Smoking, UKB = UK Biobank. * = operationalized as Chronic Ischemic Heart Disease diagnosis in replication GWAS. ** = operationalized as predicted percentage of FEV₁ in replication GWAS.

Trait	Operationalization	Variable	Discovery GWAS			Replication GWAS		r_g
			Source	Summary statistics?	<i>N</i>	Source	<i>N</i>	
ADHD	Clinical diagnosis	Binary	PGC & iPSYCH-ADHD cohort (Demontis et al., 2019)	Yes	55.374	PGC & iPSYCH-ADHD cohort (Demontis et al., 2019)	55.374	1
AU	AUDIT-C score	Categorical	Million Veteran Program (Kranzler et al., 2019)	No	274.424	UKB	147.267	x
CAD	Inclusive CAD diagnosis *	Binary	CARDIoGRAMplusC4D Consortium (Nikpay et al., 2015)	No	187.599	UKB (Neale Lab, I25)	361.194	x
EA	Years of completed schooling	Categorical	EA3 (Lee et al., 2018) excl. British cohorts and 23&Me	Yes	242.952	UKB	451.247	0.964
COPD	Inverse ratio forced expiratory volume in one second (FEV ₁) to forced vital capacity (FVC) **	Continuous	Soler Artigas et al. (2011)	No	94.612	UKB (Neale Lab, 20154)	117.241	x
SM	Smoking initiation (ever been a regular smoker yes/no)	Binary	GSCAN (Liu et al., 2019) excl. UKB and 23&Me	Yes	249.171	UKB	301.733	0.610

Table 2: Genetic (partial) correlations. For the correlations, the estimate (r_g) standard error (S.E.), an uncorrected and false discovery rate (FDR) corrected P value (P and Q respectively) were reported. For partial correlations, only the estimate (pr_g) and P value were available. * indicate a significant correlation (P or $Q < 0.05$). ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, n.s. = not significant, S.E. = standard error, SM = Smoking.

Exposure	Outcome	Correlations				Partial Correlations	
		r_g	S.E.	P	Q	pr_g	P
ADHD	AU	-0.107	0.039	0.006 *	0.038 *	-0.111	n.s.
ADHD	CAD	0.193	0.044	1.10×10^{-5} *	1.08×10^{-4} *	0.105	n.s.
ADHD	EA	-0.489	0.027	1.06×10^{-72} *	4.76×10^{-71} *	-0.412	< 0.05 *
ADHD	COPD	0.169	0.033	3.34×10^{-7} *	4.86×10^{-6} *	-0.063	< 0.05 *
ADHD	SM	0.237	0.037	1.97×10^{-10} *	3.46×10^{-9} *	0.234	< 0.05 *
AU	CAD	0.011	0.045	0.809	1	0.089	n.s.
AU	EA	0.195	0.029	3.74×10^{-11} *	7.85×10^{-10} *	0.184	< 0.05 *
AU	COPD	-0.059	0.037	0.113	0.374	0.051	n.s.
AU	SM	0.358	0.040	3.75×10^{-19} *	1.38×10^{-17} *	0.408	< 0.05 *
CAD	EA	-0.225	0.029	2.41×10^{-14} *	6.07×10^{-13} *	-0.151	< 0.05 *
CAD	COPD	0.161	0.037	1.65×10^{-5} *	1.30×10^{-4} *	-0.122	< 0.05 *
CAD	SM	0.006	0.040	0.883	1	-0.079	n.s.
EA	COPD	-0.188	0.022	7.62×10^{-18} *	2.31×10^{-16} *	0.093	< 0.05 *
EA	SM	-0.117	0.026	5.68×10^{-6} *	6.81×10^{-5} *	-0.077	n.s.
COPD	SM	0.075	0.038	0.051	0.238	-0.058	n.s.

Table 3: Total causal effects of Attention-Deficit/ Hyperactivity Disorder (ADHD), alcohol use (AU), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), educational attainment (EA), and smoking initiation (SM) on these respective traits. Results from inverse-variance weighted (IVW) method, MR-Egger regression and the weighted median and mode methods. * indicate a significant causal effect (P or $Q < 0.05$). # SNPs = the number of single nucleotide polymorphisms in the genetic instrumental variable (IV) for a specific trait are presented. Both an uncorrected and false discovery rate (FDR) corrected P value (P and Q respectively) are reported. S.E. = Standard Error.

Exposure	Outcome	Method	# SNPs	Estimate	S.E.	P	Q
ADHD	AU	IVW	11	-0.011	0.006	0.095	0.219
		MR-Egger		-0.008	0.030	0.803	0.875
		Weighted Median		-0.009	0.006	0.114	0.264
		Weighted Mode		-0.011	0.009	0.218	0.413
ADHD	CAD	IVW	11	0.002	0.003	0.501	0.657
		MR-Egger		0.007	0.012	0.517	0.738
		Weighted Median		0.005	0.003	0.068	0.169
		Weighted Mode		0.007	0.004	0.107	0.286
ADHD	EA	IVW	11	-0.428	0.119	$3.16 \times 10^{-4} *$	0.002 *
		MR-Egger		0.043	0.528	0.936	0.936
		Weighted Median		-0.374	0.087	$1.92 \times 10^{-5} *$	$1.66 \times 10^{-4} *$
		Weighted Mode		-0.265	0.170	0.119	0.286
ADHD	COPD	IVW	11	0.033	0.018	0.064	0.201
		MR-Egger		0.167	0.071	0.019 *	0.207
		Weighted Median		0.034	0.022	0.125	0.268
		Weighted Mode		0.039	0.034	0.248	0.413
ADHD	SM	IVW	11	0.018	0.005	$2.11 \times 10^{-4} *$	0.002 *
		MR-Egger		0.033	0.022	0.142	0.393
		Weighted Median		0.020	0.006	$4.30 \times 10^{-4} *$	0.003 *
		Weighted Mode		0.023	0.009	0.015 *	0.055
AU	ADHD	IVW	8	0.285	0.602	0.636	0.759
		MR-Egger		-0.664	0.819	0.418	0.738
		Weighted Median		0.148	0.710	0.835	0.917
		Weighted Mode		-0.249	0.745	0.738	0.860
AU	CAD	IVW	8	0.034	0.034	0.318	0.502
		MR-Egger		0.077	0.047	0.102	0.393
		Weighted Median		0.076	0.028	0.007 *	0.023 *
		Weighted Mode		0.072	0.027	0.007 *	0.032 *
AU	EA	IVW	8	-1.601	1.413	0.257	0.456
		MR-Egger		-1.282	2.178	0.556	0.748
		Weighted Median		-2.578	0.608	$2.22 \times 10^{-5} *$	$1.66 \times 10^{-4} *$
		Weighted Mode		-2.608	0.618	$2.42 \times 10^{-5} *$	$3.62 \times 10^{-4} *$

AU	COPD	IVW	8	0.922	0.238	$1.06 \times 10^{-4} *$	$1.06 \times 10^{-3} *$
		MR-Egger		0.863	0.373	0.021 *	0.207
		Weighted Median		0.681	0.249	0.006 *	0.023 *
		Weighted Mode		0.680	0.252	0.007 *	0.032 *
AU	SM	IVW	8	-0.012	0.044	0.777	0.832
		MR-Egger		-0.056	0.063	0.378	0.720
		Weighted Median		-0.046	0.054	0.395	0.565
		Weighted Mode		-0.048	0.054	0.377	0.565
CAD	ADHD	IVW	36	0.063	0.922	0.946	0.946
		MR-Egger		-2.564	1.600	0.109	0.393
		Weighted Median		-0.366	1.091	0.737	0.885
		Weighted Mode		-0.519	1.104	0.638	0.832
CAD	AU	IVW	36	0.131	0.072	0.067	0.201
		MR-Egger		0.064	0.129	0.622	0.778
		Weighted Median		0.104	0.089	0.243	0.384
		Weighted Mode		0.135	0.087	0.122	0.286
CAD	EA	IVW	36	0.327	0.773	0.673	0.759
		MR-Egger		-0.431	1.395	0.757	0.875
		Weighted Median		-0.028	0.864	0.974	0.974
		Weighted Mode		-0.022	0.851	0.980	0.980
CAD	COPD	IVW	36	-0.271	0.312	0.386	0.551
		MR-Egger		-0.131	0.564	0.817	0.875
		Weighted Median		0.018	0.335	0.958	0.974
		Weighted Mode		-0.011	0.337	0.973	0.980
CAD	SM	IVW	36	0.182	0.072	0.012	0.051
		MR-Egger		-0.031	0.124	0.802	0.875
		Weighted Median		0.102	0.081	0.204	0.361
		Weighted Mode		0.123	0.077	0.108	0.286
EA	ADHD	IVW	37	-0.221	0.028	$7.40 \times 10^{-15} *$	$2.22 \times 10^{-13} *$
		MR-Egger		-0.186	0.070	0.008 *	0.207
		Weighted Median		-0.218	0.045	$1.11 \times 10^{-6} *$	$3.32 \times 10^{-5} *$
		Weighted Mode		-0.242	0.076	0.002 *	0.011 *
EA	AU	IVW	37	0.006	0.004	0.128	0.275
		MR-Egger		0.010	0.009	0.255	0.546
		Weighted Median		0.007	0.004	0.047 *	0.128
		Weighted Mode		0.006	0.005	0.228	0.413
EA	CAD	IVW	37	-0.002	9.72×10^{-4}	0.088	0.219

			MR-Egger		-0.004	0.002	0.066	0.393
			Weighted Median		-0.002	0.001	0.169	0.325
			Weighted Mode		-0.003	0.002	0.124	0.286
			IVW		-0.020	0.015	0.177	0.354
EA	COPD	37	MR-Egger		-0.053	0.037	0.148	0.393
			Weighted Median		-0.031	0.015	0.037 *	0.111
			Weighted Mode		-0.047	0.017	0.006 *	0.032 *
			IVW		-0.001	0.003	0.683	0.759
			MR-Egger		-0.005	0.008	0.573	0.748
			Weighted Median		-0.002	0.003	0.479	0.653
			Weighted Mode		-0.002	0.004	0.636	0.832
			IVW		-0.120	0.266	0.651	0.759
			MR-Egger		1.147	0.541	0.034 *	0.254
			Weighted Median		0.066	0.301	0.827	0.917
			Weighted Mode		0.102	0.316	0.746	0.860
			IVW		0.016	0.017	0.338	0.507
			MR-Egger		0.061	0.040	0.126	0.393
			Weighted Median		0.026	0.021	0.234	0.384
			Weighted Mode		0.028	0.023	0.229	0.413
			IVW		0.031	0.014	0.030 *	0.112
			MR-Egger		0.024	0.036	0.497	0.738
			Weighted Median		0.018	0.013	0.173	0.325
			Weighted Mode		0.000	0.017	0.958	0.980
			IVW		0.029	0.194	0.880	0.910
			MR-Egger		0.408	0.468	0.384	0.720
			Weighted Median		-0.045	0.247	0.856	0.917
			Weighted Mode		-0.227	0.301	0.451	0.644
			IVW		-0.020	0.020	0.316	0.502
			MR-Egger		-0.068	0.048	0.157	0.393
			Weighted Median		-0.013	0.022	0.564	0.736
			Weighted Mode		-0.007	0.027	0.804	0.893
			IVW		11.982	3.679	0.001 *	0.006 *
			MR-Egger		-4.271	6.500	0.511	0.738
			Weighted Median		8.077	2.786	0.004 *	0.019 *
			Weighted Mode		7.430	2.157	5.72 x 10 ⁻⁴ *	0.006 *
			IVW		0.132	0.197	0.504	0.657
			MR-Egger		0.611	0.523	0.243	0.546

SM	CAD	6	Weighted Median	0.160	0.162	0.322	0.483
			Weighted Mode	0.156	0.158	0.324	0.512
			IVW	0.085	0.075	0.258	0.456
			MR-Egger	-0.246	0.141	0.083	0.393
			Weighted Median	0.027	0.067	0.688	0.860
			Weighted Mode	0.021	0.063	0.736	0.860
SM	EA	6	IVW	-9.810	5.778	0.090	0.219
			MR-Egger	2.691	15.777	0.865	0.894
			Weighted Median	-7.674	2.843	0.007 *	0.023 *
			Weighted Mode	-2.408	2.029	0.235	0.413
SM	COPD	6	IVW	3.697	0.752	$8.72 \times 10^{-7} *$	$1.31 \times 10^{-5} *$
			MR-Egger	1.292	1.814	0.476	0.738
			Weighted Median	3.424	0.797	$1.76 \times 10^{-5} *$	$1.66 \times 10^{-4} *$
			Weighted Mode	3.318	0.764	$1.40 \times 10^{-5} *$	$3.62 \times 10^{-4} *$

Table 4: Direct causal effects of Attention-Deficit/Hyperactivity Disorder (ADHD), alcohol use (AU), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), educational attainment (EA) and smoking initiation (SM) on these respective traits. Results from multivariable inverse-variance weighted (MV IVW) method and multivariable MR-Egger regression (MV MR-Egger) are shown. * indicate a significant causal effect (P or $Q < 0.05$). The number of single nucleotide polymorphisms (SNPs) part of the genetic instrumental variable (IV) for a specific trait are presented. Both an uncorrected and false discovery rate (FDR) corrected P value (P and Q respectively) are reported. S.E. = Standard Error.

Exposure	Outcome	Method	# SNPs	Estimate	S.E.	P	Q
ADHD	AU	MV IVW	11	-0.004	0.017	0.839	0.899
		MV MR-Egger		0.010	0.072	0.890	1
ADHD	CAD	MV IVW	11	0.002	0.006	0.779	0.899
		MV MR-Egger		0.014	0.023	0.534	1
ADHD	EA	MV IVW	11	-0.392	0.268	0.143	0.519
		MV MR-Egger		-0.394	1.281	0.758	1
ADHD	COPD	MV IVW	11	0.058	0.037	0.113	0.483
		MV MR-Egger		0.212	0.106	0.045 *	0.448
ADHD	SM	MV IVW	11	0.021	0.010	0.042 *	0.318
		MV MR-Egger		0.070	0.023	0.003 *	0.064
AU	ADHD	MV IVW	8	1.978	2.557	0.439	0.732
		MV MR-Egger		0.138	3.626	0.970	1
AU	CAD	MV IVW	8	0.112	0.135	0.408	0.721
		MV MR-Egger		0.155	0.181	0.394	0.984
AU	EA	MV IVW	8	-6.658	5.285	0.208	0.519
		MV MR-Egger		-7.462	8.070	0.355	0.969
AU	COPD	MV IVW	8	1.114	0.249	7.72×10^{-6} *	1.16×10^{-4} *
		MV MR-Egger		1.246	0.436	0.004 *	0.064
AU	SM	MV IVW	8	-0.193	0.113	0.087	0.433
		MV MR-Egger		-0.217	0.117	0.063	0.454
CAD	ADHD	MV IVW	36	0.107	1.098	0.922	0.947
		MV MR-Egger		-2.691	1.682	0.110	0.469
CAD	AU	MV IVW	36	0.099	0.075	0.186	0.519
		MV MR-Egger		0.090	0.128	0.481	1
CAD	EA	MV IVW	36	-0.448	0.866	0.605	0.757
		MV MR-Egger		-0.853	1.457	0.558	1
CAD	COPD	MV IVW	36	-0.439	0.347	0.205	0.519
		MV MR-Egger		-0.244	0.594	0.682	1
CAD	SM	MV IVW	36	0.144	0.078	0.064	0.383
		MV MR-Egger		-0.068	0.129	0.598	1
EA	ADHD	MV IVW	37	-0.209	0.032	6.69×10^{-11} *	2.01×10^{-9} *
		MV MR-Egger		-0.138	0.079	0.082	0.454

EA	AU	MV IVW	37	0.006	0.006	0.312	0.668
		MV MR-Egger		0.016	0.010	0.130	0.488
EA	CAD	MV IVW	37	-0.002	0.002	0.352	0.682
		MV MR-Egger		-0.005	0.003	0.091	0.454
EA	COPD	MV IVW	37	-0.008	0.026	0.771	0.899
		MV MR-Egger		-0.046	0.044	0.289	0.963
EA	SM	MV IVW	37	0.004	0.006	0.490	0.757
		MV MR-Egger		-0.003	0.009	0.722	1
COPD	ADHD	MV IVW	9	-0.164	0.315	0.602	0.757
		MV MR-Egger		0.666	0.709	0.347	0.969
COPD	AU	MV IVW	9	0.013	0.022	0.552	0.757
		MV MR-Egger		0.000	0.063	0.991	1
COPD	CAD	MV IVW	9	0.020	0.022	0.364	0.682
		MV MR-Egger		0.025	0.071	0.721	1
COPD	EA	MV IVW	9	-0.064	0.316	0.839	0.899
		MV MR-Egger		0.073	0.944	0.938	1
COPD	SM	MV IVW	9	-0.018	0.035	0.596	0.757
		MV MR-Egger		-0.040	0.104	0.698	1
SM	ADHD	MV IVW	6	10.748	20.166	0.594	0.757
		MV MR-Egger		-23.981	x	x	x
SM	AU	MV IVW	6	1.109	0.872	0.203	0.519
		MV MR-Egger		3.309	x	x	x
SM	CAD	MV IVW	6	-0.340	0.320	0.287	0.663
		MV MR-Egger		-0.445	x	x	x
SM	EA	MV IVW	6	2.605	38.990	0.947	0.947
		MV MR-Egger		-126.339	x	x	x
SM	COPD	MV IVW	6	3.889	1.727	0.024 *	0.243
		MV MR-Egger		12.052	x	x	x

Table 5: Summary of the hypothesized and significant correlations and causal relationships between Attention-Deficit/Hyperactivity Disorder (ADHD), alcohol use (AU), chronic obstructive pulmonary disease (COPD), coronary artery disease (CAD), educational attainment (EA) and smoking initiation (SM). DAG = directed acyclic graph, r_g = correlation estimate, pr_g = partial correlation estimate, IVW = inverse-variance weighted method, MVIVW = multivariable inverse-variance weighted method, + = positive relationship, (+) = positive relationship only when not correcting for multiple testing (P , only for IVW and MVIVW), - = negative relationship, (-) = negative relationship only when not correcting for multiple testing (P , only for IVW and MVIVW), x = relationship whose direction was not specified.

Correlation			Hypothesis	r_g	pr_g	Causation		Hypothesis	IVW	MVIVW	DAG
ADHD	≈	AU	+	-		ADHD	→	AU	+		
						AU	→	ADHD			
ADHD	≈	CAD		+		ADHD	→	CAD			
						CAD	→	ADHD			
ADHD	≈	EA	-	-	-	ADHD	→	EA	-	-	
						EA	→	ADHD			
ADHD	≈	COPD		+	+	ADHD	→	COPD			x
						COPD	→	ADHD			
ADHD	≈	SM	+	+	+	ADHD	→	SM	+	(+)	x
						SM	→	ADHD			
AU	≈	CAD	+			AU	→	CAD	+		
						CAD	→	AU			
AU	≈	EA	-	+	+	AU	→	EA			x
						EA	→	AU			
AU	≈	COPD				AU	→	COPD	+	+	
						COPD	→	AU			
AU	≈	SM	+	+	+	AU	→	SM	+		x
						SM	→	AU			
CAD	≈	EA		-	-	CAD	→	EA			x
						EA	→	CAD			
CAD	≈	COPD		+	+	CAD	→	COPD			
						COPD	→	CAD			
CAD	≈	SM	+			CAD	→	SM	(+)	(+)	
						SM	→	CAD			
EA	≈	COPD		-	-	EA	→	COPD			x
						COPD	→	EA			
EA	≈	SM	-	-		EA	→	SM	-		
						SM	→	EA			
COPD	≈	SM	+			COPD	→	SM			
						SM	→	COPD			

References

- Acosta MC, Eissenberg T, Nichter Mimi, Nichter Mark, Balster RL, (TERN) TTERN. 2008. Characterizing Early Cigarette Use Episodes in Novice Smokers. *Addict Behav* **33**:106–121. doi:10.1016/j.addbeh.2007.09.005
- Allman PH, Aban IB, Tiwari HK, Cutter GR. 2018. An Introduction to Mendelian Randomization with Applications in Neurology. *Mult Scler Relat Disord* **24**:72–78. doi:10.1016/j.msard.2018.06.017
- American Psychological Association. 2013. Diagnostic and Statistical Manual of Mental Disorders (DSM-V), 5th ed. Washington, DC: American Psychiatric Publishing.
- Angrist JD, Imbens GW, Rubin DB. 1996. Identification of Causal Effects Using Instrumental Variables. *J Am Stat Assoc* **91**:444–455. doi:10.1080/01621459.1996.10476902
- Au Yeung SL, Borges MC, Lawlor DA. 2018. Association of Genetic Instrumental Variables for Lung Function on Coronary Artery Disease Risk: A 2-Sample Mendelian Randomization Study. *Circ Genomic Precis Med* **11**:e001952. doi:10.1161/CIRCGEN.117.001952
- Barbor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. 2001. Alcohol Use Disorders Identification Test (AUDIT) Guidelines for Use in Primary Care, Second. ed. Geneva. doi:10.1177/0269881110393051
- Bobo JK, Husten C. 2000. Sociocultural Influences on Smoking and Drinking. *Alcohol Res Heal* **24**:225–232.
- Boef AGC, Dekkers OM, Le Cessie S. 2015. Mendelian randomization studies: A Review of the Approaches Used and the Quality of Reporting. *Int J Epidemiol* **44**:496–511. doi:10.1093/ije/dyv071
- Bowden J, Davey Smith G, Burgess S. 2015. Mendelian Randomization With Invalid Instruments: Effect Estimation and Bias Detection Through Egger Regression. *Int J Epidemiol* **44**:512–525. doi:10.1093/ije/dyv080
- Bowden J, Davey Smith G, Haycock PC, Burgess S. 2016. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* **40**:304–314. doi:10.1002/gepi.21965
- Bowden J, Del Greco M F, Minelli C, Zhao Q, Lawlor DA, Sheehan NA, Thompson J, Davey Smith G. 2019. Improving the Accuracy of Two-Sample Summary-Data Mendelian Randomization: Moving Beyond the NOME Assumption. *Int J Epidemiol* **48**:728–742. doi:10.1093/ije/dyy258
- Bowden J, Dudbridge F. 2009. Unbiased Estimation of Odds Ratios: Combining Genomewide Association Scans with Replication Studies. *Genet Epidemiol* **33**:406–418. doi:10.1002/gepi.20394
- Brinkman WB, Epstein JN, Auinger P, Tamm L, Froehlich TE. 2015. Association of Attention-

- Deficit/Hyperactivity Disorder and Conduct Disorder with Early Tobacco and Alcohol Use. *Drug Alcohol Depend* **147**:183–189. doi:10.1016/j.drugalcdep.2014.11.018
- Brumpton B, Sanderson E, Hartwig FP, Harrison S, Vie GÅ, Cho Y, Howe LD, Hughes A, Boomsma DI, Havdahl A, Hopper J, Neale M, Nivard MG, Pedersen NL, Reynolds CA, Tucker-Drob EM, Grotzinger A, Howe L, Morris T, Li S, Consortium M within-family, Chen W-M, Bjørngaard JH, Hveem K, Willer C, Evans DM, Kaprio J, Davey Smith G, Åsvold BO, Hemani G, Davies NM. 2019. Within-Family Studies for Mendelian Randomization: Avoiding Dynastic, Assortative Mating, and Population Stratification Biases. *bioRxiv* 602516. doi:10.1101/602516
- Burgess S, Davies NM, Thompson SG. 2016. Bias Due to Participant Overlap in Two-Sample Mendelian Randomization. *Genet Epidemiol* **40**:597–608. doi:10.1002/gepi.21998
- Burgess S, Thompson SG. 2015. Multivariable Mendelian Randomization: The Use of Pleiotropic Genetic Variants to Estimate Causal Effects. *Am J Epidemiol* **181**:251–260. doi:10.1093/aje/kwu283
- Burgess S, Thompson SG. 2011. Avoiding Bias from Weak Instruments in Mendelian Randomization Studies. *Int J Epidemiol* **40**:755–764. doi:10.1093/ije/dyr036
- Carter P, Lagan J, Fortune C, Bhatt DL, Vestbo J, Niven R, Chaudhuri N, Schelbert EB, Potluri R, Miller CA. 2019. Association of Cardiovascular Disease With Respiratory Disease. *J Am Coll Cardiol* **73**:2166–2177. doi:10.1016/j.jacc.2018.11.063
- Charach A, Yeung E, Climans T, Lillie E. 2011. Childhood Attention-Deficit/Hyperactivity Disorder and Future Substance Use Disorders: Comparative Meta-Analyses. *J Am Acad Child Adolesc Psychiatry* **50**:9–21. doi:10.1016/j.jaac.2010.09.019
- Cochran WG. 1954. The Combination of Estimates from Different Experiments. *Biometrics* **10**:101–129.
- Colombo D, Maathuis MH. 2014. Order-Independent Causal Structure Learning Order-Independent Constraint-Based Causal Structure Learning. *J Mach Learn Res* **15**:3921–3962.
- Dai J, Mukamal KJ, Krasnow RE, Swan GE, Reed T. 2015. Higher Usual Alcohol Consumption Was Associated with a Lower 41-Y Mortality Risk from Coronary Artery Disease in Men Independent of Genetic and Common Environmental Factors: The Prospective NHLBI Twin Study. *Am J Clin Nutr* **102**:31–39. doi:10.3945/ajcn.114.106435
- Davey Smith G, Ebrahim S. 2003. “Mendelian Randomization”: Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease? *Int J Epidemiol* **32**:1–22. doi:10.1093/ije/dyg070
- Davey Smith G, Hemani G. 2014. Mendelian Randomization: Genetic Anchors for Causal Inference in Epidemiological Studies. *Hum Mol Genet* **23**:89–98. doi:10.1093/hmg/ddu328

- Davey Smith G, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. 2007. Clustered Environments and Randomized Genes: a Fundamental Distinction Between Conventional and Genetic Epidemiology. *PLoS Med* **4**:1985–1992. doi:10.1371/journal.pmed.0040352
- Del Greco M F, Minelli C, Sheehan NA, Thompson JR. 2015. Detecting Pleiotropy in Mendelian Randomisation Studies With Summary Data and a Continuous Outcome. *Stat Med* **34**:2926–2940. doi:10.1002/sim.6522
- Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, Baldursson G, Belliveau R, Bybjerg-Grauholm J, Bækvad-Hansen M, Cerrato F, Chambert K, Churchhouse C, Dumont A, Eriksson N, Gandal M, Goldstein JL, Grasby KL, Grove J, Gudmundsson OO, Hansen CS, Hauberg ME, Hollegaard M V., Howrigan DP, Huang H, Maller JB, Martin AR, Martin NG, Moran J, Pallesen J, Palmer DS, Pedersen CB, Pedersen MG, Poterba T, Poulsen JB, Ripke S, Robinson EB, Satterstrom FK, Stefansson H, Stevens C, Turley P, Walters GB, Won H, Wright MJ, Albayrak Ö, Anney RJL, Arranz MJ, Banaschewski TJ, Bau C, Biederman J, Buitelaar JK, Casas M, Charach A, Crosbie J, Dempfle A, Doyle AE, Ebstein RP, Elia J, Freitag C, Föcker M, Gill M, Grevet E, Hawi Z, Hebebrand J, Herpertz-Dahlmann B, Hervas A, Hinney A, Hohmann S, Holmans P, Hutz M, Ickowitz A, Johansson S, Kent L, Kittel-Schneider S, Lambregts-Rommelse N, Lehmkuhl G, Loo SK, McGough JJ, Meyer J, Mick E, Middleton F, Miranda A, Mota NR, Mulas F, Mulligan A, Nelson F, Nguyen TT, Oades RD, O'Donovan MC, Owen MJ, Palmason H, Ramos-Quiroga JA, Renner TJ, Ribasés M, Rietschel M, Rivero O, Romanos J, Romanos M, Rothenberger A, Royers H, Sánchez-Mora C, Scherag A, Schimmelmann BG, Schäfer H, Sergeant J, Sinzig J, Smalley SL, Steinhausen HC, Thompson M, Todorov A, Vasquez AA, Walitza S, Wang Y, Warnke A, Williams N, Witt SH, Yang L, Zayats T, Zhang-James Y, Smith GD, Davies GE, Ehli EA, Evans DM, Fedko IO, Greven CU, Groen-Blokhuis MM, Guxens M, Hammerschlag AR, Hartman CA, Heinrich J, Jan Hottenga J, Hudziak J, Jugessur A, Kemp JP, Krapohl E, Murcia M, Myhre R, Nolte IM, Nyholt DR, Ormel J, Ouwers KG, Pappa I, Pennell CE, Plomin R, Ring S, Standl M, Stergiakouli E, Pourcain BS, Stoltenberg C, Sunyer J, Thiering E, Tiemeier H, Tiesler CMT, Timpson NJ, Trzaskowski M, van der Most PJ, Vilor-Tejedor N, Wang CA, Whitehouse AJO, Zhao H, Agee M, Alipanahi B, Auton A, Bell RK, Bryc K, Elson SL, Fontanillas P, Furlotte NA, Hinds DA, Hromatka BS, Huber KE, Kleinman A, Litterman NK, McIntyre MH, Mountain JL, Northover CAM, Pitts SJ, Sathirapongsasuti JF, Sazonova O V., Shelton JF, Shringarpure S, Tian C, Vacic V, Wilson CH, Andreassen OA, Asherson P, Burton CL, Boomsma DI, Cormand B, Dalsgaard S, Franke B, Gelernter J, Geschwind D, Hakonarson H, Haavik J, Kranzler HR, Kuntsi J, Langley K, Lesch KP, Middeldorp C, Reif A, Rohde LA, Roussos P, Schachar R, Sklar P, Sonuga-Barke EJS, Sullivan PF, Thapar A, Tung JY, Waldman ID, Medland SE, Stefansson K, Nordentoft M, Hougaard DM, Werge T, Mors O, Mortensen PB, Daly MJ, Faraone S V., Børglum AD, Neale BM. 2019. Discovery of the First Genome-Wide Significant Risk Loci for Attention Deficit/Hyperactivity Disorder. *Nat Genet* **51**:63–75. doi:10.1038/s41588-018-0269-7
- Devine JF. 2008. Chronic Obstructive Pulmonary Disease: An Overview. *Am Heal Drug Benefits* **1**:34–42. doi:10.12968/bjon.2016.25.7.360
- Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. 2012. qgraph: Network Visualizations of Relationships in Psychometric Data. *J Stat Softw* **48**. doi:10.18637/jss.v048.i04

- Gage SH, Bowden J, Davey Smith G, Munafo MR. 2018. Investigating Causality in Associations between Education and Smoking: A Two-Sample Mendelian Randomization Study. *Int J Epidemiol* **47**:1131–1140. doi:10.1093/ije/dyy131
- Gjerdevik M, Grydeland TB, Washko GR, Coxson HO, Silverman EK, Gulsvik A, Bakke PS. 2015. The Relationship of Educational Attainment with Pulmonary Emphysema and Airway Wall Thickness. *Ann Am Thorac Soc* **12**:813–820. doi:10.1513/AnnalsATS.201410-485OC
- Greene CC, Bradley KA, Bryson CL, Blough DK, Evans LE, Udris EM, Au DH. 2008. The Association between Alcohol Consumption and Risk of COPD Exacerbation in a Veteran Population. *Chest* **134**:761–767. doi:10.1378/chest.07-3081
- Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, Ip HF, Marioni RE, McIntosh AM, Deary IJ, Koellinger PD, Harden KP, Nivard MG, Tucker-Drob EM. 2019. Genomic Structural Equation Modelling Provides Insights into the Multivariate Genetic Architecture of Complex Traits. *Nat Hum Behav* **3**:513–525. doi:10.1038/s41562-019-0566-x
- Hackshaw A, Morris JK, Boniface S, Tang JL, Milenkovi D. 2018. Low Cigarette Consumption and Risk of Coronary Heart Disease and Stroke: Meta-Analysis of 141 Cohort Studies in 55 Study Reports. *BMJ* **360**:j5855. doi:10.1136/bmj.j5855
- Harrison ELR, Desai RA, McKee SA. 2008. Nondaily Smoking and Alcohol Use, Hazardous Drinking, and Alcohol Diagnoses among Young Adults: Findings from the NESARC. *Alcohol Clin Exp Res* **32**:2081–2087. doi:10.1111/j.1530-0277.2008.00796.x
- Hartwig FP, Davey Smith G, Bowden J. 2017. Robust Inference in Summary Data Mendelian Randomization via the Zero Modal Pleiotropy Assumption. *Int J Epidemiol* **46**:1985–1998. doi:10.1093/ije/dyx102
- Hartwig FP, Davies NM, Hemani G, Davey Smith G. 2016. Counterfactual Causation: Avoiding the Downsides of a Powerful, Widely Applicable but Potentially Fallible Technique. *Int J Epidemiol* **45**:1717–1726. doi:10.1093/ije/dyx028
- Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. 2016. Best (but Oft-Forgotten) Practices: the Design, Analysis, and Interpretation of Mendelian Randomization Studies. *Am J Clin Nutr* **103**:965–78. doi:10.3945/ajcn.115.118216.INTRODUCTION
- Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM, Relton C, Martin RM, Davey Smith G, Gaunt TR, Haycock PC. 2018. The MR-Base Platform Supports Systematic Causal Inference Across the Human Phenome. *Elife* **7**:e34408. doi:10.7554/eLife.34408
- Howey R, Shin SY, Relton C, Davey Smith G, Cordell HJ. 2020. Bayesian Network Analysis Incorporating Genetic Anchors Complements Conventional Mendelian Randomization Approaches for Exploratory Analysis of Causal Relationships in Complex Data. *PLoS Genet* **16**:e1008198. doi:10.1371/journal.pgen.1008198

- Johnson JD, Theurer WM. 2014. A Stepwise Approach to the Interpretation of Pulmonary Function Tests. *Am Fam Physician* **89**:359–366.
- Johnson T. 2012. Efficient Calculation for Multi-SNP Genetic Risk Scores American Society of Human Genetics Annual Meeting. San Francisco. doi:10.1038/ng.784
- Kang H, Zhang A, Cai TT, Small DS. 2016. Instrumental Variables Estimation With Some Invalid Instruments and its Application to Mendelian Randomization. *J Am Stat Assoc* **111**:132–144. doi:10.1080/01621459.2014.994705
- Kollins SH, McClernon FJ, Fuemmeler BF. 2005. Association between Smoking and Attention-Deficit/Hyperactivity Disorder Symptoms in a Population-Based Sample of Young Adults. *Arch Gen Psychiatry* **62**:1142–1147. doi:10.1001/archpsyc.62.10.1142
- Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, Damrauer S, Tsao PS, Klarin D, Baras A, Reid J, Overton J, Rader DJ, Cheng Z, Tate JP, Becker WC, Concato J, Xu K, Polimanti R, Zhao H, Gelernter J. 2019. Genome-Wide Association Study of Alcohol Consumption and Use Disorder in 274,424 Individuals from Multiple Populations. *Nat Commun* **10**:1499. doi:10.1038/s41467-019-09480-8
- Kubota Y, Heiss G, Maclehose RF, Roetker NS, Folsom AR. 2017. Association of Educational Attainment With Lifetime Risk of Cardiovascular Disease - The Atherosclerosis Risk in Communities Study. *JAMA Intern Med* **177**:1165–1172. doi:10.1001/jamainternmed.2017.1877
- Latvala A, Rose RJ, Pulkkinen L, Dick DM, Korhonen T, Kaprio J. 2014. Drinking, Smoking, and Educational Achievement: Cross-Lagged Associations from Adolescence to Adulthood. *Drug Alcohol Depend* **137**:106–113. doi:10.1016/j.drugalcdep.2014.01.016
- Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. 2008. Mendelian Randomization: Using Genes as Instruments for Making Causal Inferences in Epidemiology. *Stat Med* **27**:1133–1163. doi:10.1002/sim
- Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, Nguyen-Viet TA, Bowers P, Sidorenko J, Karlsson Linnér R, Fontana MA, Kundu T, Lee C, Li H, Li R, Royer R, Timshel PN, Walters RK, Willoughby EA, Yengo L, Agee M, Alipanahi B, Auton A, Bell RK, Bryc K, Elson SL, Fontanillas P, Hinds DA, McCreight JC, Huber KE, Litterman NK, McIntyre MH, Mountain JL, Noblin ES, Northover CAM, Pitts SJ, Sathirapongsasuti JF, Sazonova O V., Shelton JF, Shringarpure S, Tian C, Vacic V, Wilson CH, Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, Turley P, Chen GB, Emilsson V, Meddens SFW, Oskarsson S, Pickrell JK, Thom K, Timshel P, Vlaming R de, Abdellaoui A, Ahluwalia TS, Bacelis J, Baumbach C, Bjornsdottir G, Brandsma JH, Concas MP, Derringer J, Furlotte NA, Galesloot TE, Girotto G, Gupta R, Hall LM, Harris SE, Hofer E, Horikoshi M, Huffman JE, Kaasik K, Kalafati IP, Karlsson R, Kong A, Lahti J, van der Lee SJ, Leeuw C de, Lind PA, Lindgren KO, Liu T, Mangino M, Marten J, Mihailov E, Miller MB, van der Most PJ, Oldmeadow C, Payton A, Pervjakova N, Peyrot WJ, Qian Y, Raitakari O, Rueedi R, Salvi E, Schmidt B, Schraut KE, Shi J, Smith A V., Poot RA, St Pourcain B, Teumer A, Thorleifsson G, Verweij N, Vuckovic D, Wellmann J, Westra HJ, Yang Jingyun, Zhao W, Zhu Z, Alizadeh BZ, Amin N, Bakshi A, Baumeister SE, Biino G, Bønnelykke K, Boyle PA, Campbell H, Cappuccio FP,

Davies G, De Neve JE, Deloukas P, Demuth I, Ding J, Eibich P, Eisele L, Eklund N, Evans DM, Faul JD, Feitosa MF, Forstner AJ, Gandin I, Gunnarsson B, Halldórsson B V., Harris TB, Heath AC, Hocking LJ, Holliday EG, Homuth G, Horan MA, Hottenga JJ, de Jager PL, Joshi PK, Jugessur A, Kaakinen MA, Kähönen M, Kanoni S, Keltigangas-Järvinen L, Kiemeny LALM, Kolcic I, Koskinen S, Kraja AT, Kroh M, Kutalik Z, Latvala A, Launer LJ, Lebreton MP, Levinson DF, Lichtenstein P, Lichtner P, Liewald DCM, Loukola, Life Lines Cohort Study A, Madden PA, Mägi R, Mäki-Opas T, Marioni RE, Marques-Vidal P, Meddens GA, McMahon G, Meisinger C, Meitinger T, Milaneschi Y, Milani L, Montgomery GW, Myhre R, Nelson CP, Nyholt DR, Ollier WER, Palotie A, Paternoster L, Pedersen NL, Petrovic KE, Porteous DJ, Räikkönen K, Ring SM, Robino A, Rostapshova O, Rudan I, Rustichini A, Salomaa V, Sanders AR, Sarin AP, Schmidt H, Scott RJ, Smith BH, Smith JA, Staessen JA, Steinhagen-Thiessen E, Strauch K, Terracciano A, Tobin MD, Ulivi S, Vaccargiu S, Quaye L, van Rooij FJA, Venturini C, Vinkhuyzen AAE, Völker U, Völzke H, Vonk JM, Vozzi D, Waage J, Ware EB, Willemsen G, Attia JR, Bennett DA, Berger K, Bertram L, Bisgaard H, Boomsma DI, Borecki IB, Bültmann U, Chabris CF, Cucca F, Cusi D, Deary IJ, Dedoussis G V., van Duijn CM, Eriksson JG, Franke B, Franke L, Gasparini P, Gejman P V., Gieger C, Grabe HJ, Gratten J, Groenen PJF, Gudnason V, van der Harst P, Hayward C, Hinds DA, Hoffmann W, Hyppönen E, Iacono WG, Jacobsson B, Järvelin MR, Jöckel KH, Kaprio J, Kardina SLR, Lehtimäki T, Lehrer SF, Magnusson PKE, Martin NG, McGue M, Metspalu A, Pendleton N, Penninx BWJH, Perola M, Pirastu N, Pirastu M, Polasek O, Posthuma D, Power C, Province MA, Samani NJ, Schlessinger D, Schmidt R, Sørensen TIA, Spector TD, Stefansson K, Thorsteinsdottir U, Thurik AR, Timpson NJ, Tiemeier H, Tung JY, Uitterlinden AG, Vitart V, Vollenweider P, Weir DR, Wilson JF, Wright AF, Conley DC, Krueger RF, Smith GD, Hofman A, Laibson DI, Medland SE, Meyer MN, Yang Jian, Johannesson M, Visscher PM, Esko T, Koellinger PD, Cesarini D, Benjamin DJ, Alver M, Bao Y, Clark DW, Day FR, Furlotte NA, Joshi PK, Kemper KE, Kleinman A, Langenberg C, Mägi R, Trampush JW, Verma SS, Wu Y, Lam M, Zhao JH, Zheng Z, Boardman JD, Campbell H, Freese J, Harris KM, Hayward C, Herd P, Kumari M, Lencz T, Luan J, Malhotra AK, Metspalu A, Milani L, Ong KK, Perry JRB, Porteous DJ, Ritchie MD, Smart MC, Smith BH, Tung JY, Wareham NJ, Wilson JF, Beauchamp JP, Conley DC, Esko T, Lehrer SF, Magnusson PKE, Oskarsson S, Pers TH, Robinson MR, Thom K, Watson C, Chabris CF, Meyer MN, Laibson DI, Yang Jian, Johannesson M, Koellinger PD, Turley P, Visscher PM, Benjamin DJ, Cesarini D. 2018. Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals. *Nat Genet* **50**:1112–1121. doi:10.1038/s41588-018-0147-3

Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, Datta G, Davila-Velderrain J, McGuire D, Tian C, Zhan X, Agee M, Alipanahi B, Auton A, Bell RK, Bryc K, Elson SL, Fontanillas P, Furlotte NA, Hinds DA, Hromatka BS, Huber KE, Kleinman A, Litterman NK, McIntyre MH, Mountain JL, Northover CAM, Sathirapongsasuti JF, Sazonova O V., Shelton JF, Shringarpure S, Tung JY, Vacic V, Wilson CH, Pitts SJ, Mitchell A, Skogholt AH, Winsvold BS, Sivertsen B, Stordal E, Morken G, Kallestad H, Heuch I, Zwart JA, Fjukstad KK, Pedersen LM, Gabrielsen ME, Johnsen MB, Skrove M, Indredavik MS, Drange OK, Bjerkeset O, Børte S, Stensland SØ, Choquet H, Docherty AR, Faul JD, Foerster JR, Fritsche LG, Gordon SD, Haessler J, Hottenga JJ, Huang H, Jang SK, Jansen PR, Ling Y, Mägi R, Matoba N, McMahon G, Mulas A, Orrù V, Palviainen T, Pandit A, Reginsson GW, Smith JA, Taylor AE, Turman C, Willemsen G, Young H, Young KA, Zajac GJM, Zhao W, Zhou W, Björnsdottir G, Boardman

- JD, Boehnke M, Boomsma DI, Chen C, Cucca F, Davies GE, Eaton CB, Ehringer MA, Esko T, Fiorillo E, Gillespie NA, Gudbjartsson DF, Haller T, Harris KM, Heath AC, Hewitt JK, Hickie IB, Hokanson JE, Hopfer CJ, Hunter DJ, Iacono WG, Johnson EO, Kamatani Y, Kardia SLR, Keller MC, Kellis M, Kooperberg C, Kraft P, Krauter KS, Laakso M, Lind PA, Loukola A, Lutz SM, Madden PAF, Martin NG, McGue M, McQueen MB, Medland SE, Metspalu A, Mohlke KL, Nielsen JB, Okada Y, Peters U, Polderman TJC, Posthuma D, Reiner AP, Rice JP, Rimm E, Rose RJ, Runarsdottir V, Stallings MC, Stančáková A, Stefansson H, Thai KK, Tindle HA, Tyrfingsson T, Wall TL, Weir DR, Weisner C, Whitfield JB, Yin J, Zuccolo L, Bierut LJ, Hveem K, Lee JJ, Munafò MR, Saccone NL, Willer CJ, Cornelis MC, David SP, Jorgenson E, Kaprio J, Stitzel JA, Stefansson K, Thorgeirsson TE, Abecasis G, Liu DJ, Vrieze S. 2019. Association Studies of up to 1.2 Million Individuals Yield New Insights Into the Genetic Etiology of Tobacco and Alcohol Use. *Nat Genet* **51**:237–244. doi:10.1038/s41588-018-0307-5
- Machiela MJ, Chanock SJ. 2015. LDlink: A Web-Based Application for Exploring Population-Specific Haplotype Structure and Linking Correlated Alleles of Possible Functional Variants. *Bioinformatics* **31**:3555–3557. doi:10.1093/bioinformatics/btv402
- McKee SA, Falba T, O'Malley SS, Sindelar J, O'Connor PG. 2007. Smoking Status as a Clinical Indicator for Alcohol Misuse in US Adults. *Arch Intern Med* **167**:716–721. doi:10.1001/archinte.167.7.716
- McKee SA, Hinson R, Rounsaville D, Petrelli P. 2004. Survey of Subjective Effects of Smoking while Drinking among College Students. *Nicotine Tob Res* **6**:111–117. doi:10.1080/14622200310001656939
- Milberger S, Biederman J, Faraone S V., Chen L, Jones J. 1997. ADHD is Associated with Early Initiation of Cigarette Smoking in Children and Adolescents. *J Am Acad Child Adolesc Psychiatry* **36**:37–44. doi:10.1097/00004583-199701000-00015
- Millwood IY, Walters RG, Mei XW, Guo Y, Yang L, Bian Z, Bennett DA, Chen Y, Dong C, Hu R, Zhou G, Yu B, Jia W, Parish S, Clarke R, Davey Smith G, Collins R, Holmes M V., Li L, Peto R, Chen Z. 2019. Conventional and Genetic Evidence on Alcohol and Vascular Disease Aetiology: A Prospective Study of 500 000 Men and Women in China. *Lancet* **393**:1831–1842. doi:10.1016/S0140-6736(18)31772-0
- Moffitt RA. 1996. Identification of Causal Effects Using Instrumental Variables - Comment. *J Am Stat Assoc* **91**:462–465.
- Momen NC, Plana-Ripoll O, Agerbo E, Benros ME, Borglum AD, Christensen MK, Dalsgaard S, Degenhardt L, de Jonge P, Debois J-CPG, Fenger-Gron M, Gunn JM, Iburg KM, Kessing LV, Kessler RC, Laursen TM, Lim CCW, Mors O, Mortensen PB, MUSliner KL, Nordentoft M, Pedersen CB, Petersen LV, Ribe AR, Roest AM, Saha S, Schork AJ, Scott KM, Sievert C, Sorensen HJ, Stedman TJ, Vestergaard M, Vilhjalmsen B, Werge T, Weyerer N, Whiteford HA, Prior A, McGrath JJ. 2020. Association between Mental Disorders and Subsequent Medical Conditions. *N Engl J Med* **382**:1721–1731. doi:10.1056/NEJMoa1915784
- Mostofsky E, Chahal HS, Mukamal KJ, Rimm EB, Mittleman MA. 2016. Alcohol and Immediate Risk of Cardiovascular Events: A Systematic Review and Dose-Response Meta-Analysis. *Circulation* **133**:979–987. doi:10.1016/j.physbeh.2017.03.040

- Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, CHopewell J, Webb TR, Zeng L, Dehghan A, Alver M, MArmasu S, Auro K, Bjorres A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang Jie, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikäinen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith A V., Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, De Andrade M, De Vries PS, Van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang Jianfeng, Jalilzadeh S, Kessler T, König IR, Lannfelt L, Lieb W, Lind L, MLindgren C, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FUR, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, ESTirups K, Trompet S, Wang L, Zaman KS, Ardisino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples L, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, SHall A, Hamsten A, Harris TB, LHazen S, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimäki T, Loos RJF, Melander O, Metspalu A, März W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M. 2015. A Comprehensive 1000 Genomes-Based Genome-Wide Association Meta-Analysis of Coronary Artery Disease. *Nat Genet* **47**:1121–1130. doi:10.1038/ng.3396
- Nowak C. 2018. Lung Function and Coronary Artery Disease Risk. *Circ Genomic Precis Med* **11**:e002137. doi:10.1161/CIRCGEN.118.002137
- Osland S, Hirsch L, Pringsheim T. 2017. Smoking, Alcohol and Drug Use in Youth and Adults with Attention-Deficit Hyperactivity Disorder. *BJPsych Open* **3**:141–146. doi:10.1192/bjpo.bp.116.004317
- Pauwels R, Buist S, Calverley P, Jenkins C, Hurd S. 2001. Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Pulmonary Disease. NHBLI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop Summary. *Am J Respir Crit Care Med* **163**:1256–1276. doi:10.1016/S0873-2159(15)30846-1
- Polanczyk G V., Salum GA, Sugaya LS, Caye A, Rohde LA. 2015. Annual Research Review: A Meta-Analysis of the Worldwide Prevalence of Mental Disorders in Children and Adolescents. *J Child Psychol Psychiatry* **56**:345–365. doi:10.1111/jcpp.12381
- Polderman TJC, Boomsma DI, Bartels M, Verhulst FC, Huizink AC. 2010. A Systematic Review of Prospective Studies on Attention Problems and Academic Achievement. *Acta Psychiatr Scand* **122**:271–284. doi:10.1111/j.1600-0447.2010.01568.x
- Pomerleau OF, Downey KK, Stelson FW, Pomerleau CS. 1995. Cigarette Smoking in Adult Patients Diagnosed with Attention Deficit Hyperactivity Disorder. *J Subst Abuse* **7**:373–378. doi:10.1016/0899-3289(95)90030-6
- Qureshi AI, Suri MFK, Saad M, Hopkins LN. 2003. Educational Attainment and Risk of Stroke and Myocardial Infarction. *Med Sci Monit* **9**:466–474.

- Rees JMB, Wood AM, Burgess S. 2017. Extending the MR-Egger Method for Multivariable Mendelian Randomization to Correct for Both Measured and Unmeasured Pleiotropy. *Stat Med* **36**:4705–4718. doi:10.1002/sim.7492
- Roerecke M, Greenfield TK, Kerr WC, Bondy S, Cohen J, Rehm J. 2011. Heavy Drinking Occasions in Relation to Ischaemic Heart Disease Mortality - An 11-22 Year Follow-Up of the 1984 and 1995 US National Alcohol Surveys. *Int J Epidemiol* **40**:1401–1410. doi:10.1093/ije/dyr129
- Roerecke M, Rehm J. 2010. Irregular Heavy Drinking Occasions and Risk of Ischemic Heart Disease: A Systematic Review and Meta-Analysis. *Am J Epidemiol* **171**:633–644. doi:10.1093/aje/kwp451
- Rohrer JM. 2018. Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Adv Methods Pract Psychol Sci* **1**:27–42. doi:10.1177/2515245917745629
- Rosoff DB, Clarke TK, Adams MJ, McIntosh AM, Davey Smith G, Jung J, Lohoff FW. 2019. Educational Attainment Impacts Drinking Behaviors and Risk for Alcohol Dependence: Results from a Two-Sample Mendelian Randomization Study with ~780,000 Participants. *Mol Psychiatry*. doi:10.1038/s41380-019-0535-9
- Sanderson E, Davey Smith G, Windmeijer F, Bowden J. 2019. An Examination of Multivariable Mendelian Randomization in the Single-Sample and Two-Sample Summary Data Settings. *Int J Epidemiol* **48**:713–727. doi:10.1093/ije/dyy262
- Sanderson E, Spiller W, Bowden J. 2020. Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomisation. *bioRxiv*. doi:10.1101/2020.04.02.021980
- Saunders JB, Aasland OG, Babor TF, De La Fuente JR, Grant M. 1993. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction* **88**:791–804. doi:10.1111/j.1360-0443.1993.tb02093.x
- Schorling JB, Gutgesell M, Klas P, Smith D, Keller A. 1994. Tobacco, Alcohol and Other Drug Use among College Students. *J Subst Abuse* **6**:105–115. doi:10.1016/S0899-3289(94)90143-0
- Semeijn EJ, Sandra Kooij JJ, Comijs HC, Michielsen M, Deeg DJH, Beekman ATF. 2013. Attention-Deficit/Hyperactivity Disorder, Physical Health, and Lifestyle in Older Adults. *J Am Geriatr Soc* **61**:882–887. doi:10.1111/jgs.12261
- Simon V, Czobor P, Bálint S, Mészáros Á, Bitter I. 2009. Prevalence and Correlates of Adult Attention-Deficit Hyperactivity Disorder: Meta-Analysis. *Br J Psychiatry* **194**:204–211. doi:10.1192/bjp.bp.107.048827
- Siu ST, Udaltsova N, Iribarren C, Klatsky AL. 2010. Alcohol and Lung Airways Function. *Perm J* **14**:11–18. doi:10.7812/tpp/09-089

- Soler Artigas M, Loth DW, Wain L V., Gharib SA, Obeidat M, Tang W, Zhai G, Zhao JH, Smith AV, Huffman JE, Albrecht E, Jackson CM, Evans DM, Cadby G, Fornage M, Manichaikul A, Lopez LM, Johnson T, Aldrich MC, Aspelund T, Barroso I, Campbell H, Cassano PA, Couper DJ, Eiriksdottir G, Franceschini N, Garcia M, Gieger C, Gislason GK, Grkovic I, Hammond CJ, Hancock DB, Harris TB, Ramasamy A, Heckbert SR, Heliövaara M, Homuth G, Hysi PG, James AL, Jankovic S, Joubert BR, Karrasch S, Klopp N, Koch B, Kritchevsky SB, Launer LJ, Liu Y, Loehr LR, Lohman K, Loos RJF, Lumley T, Al Balushi KA, Ang WQ, Barr RG, Beilby J, Blakey JD, Boban M, Boraska V, Brisman J, Britton JR, Brusselle GG, Cooper C, Curjuric I, Dahgam S, Deary IJ, Ebrahim S, Eijgelsheim M, Francks C, Gaysina D, Granel R, Gu X, Hankinson JL, Hardy R, Harris SE, Henderson J, Henry A, Hingorani AD, Hofman A, Holt PG, Hui J, Hunter ML, Imboden M, Jameson KA, Kerr SM, Kolcic I, Kronenberg F, Liu JZ, Marchini J, McKeever T, Morris AD, Olin AC, Porteous DJ, Postma DS, Rich SS, Ring SM, Rivadeneira F, RoCHAT T, Sayer AA, Sayers I, Sly PD, Smith GD, Sood A, Starr JM, Uitterlinden AG, Vonk JM, Wannamethee SG, Whincup PH, Wijmenga C, Williams OD, Wong A, Mangino M, Marciante KD, McArdle WL, Meibohm B, Morrison AC, North KE, Omenaas E, Palmer LJ, Pietiläinen KH, Pin I, Polašek O, Pouta A, Psaty BM, Hartikainen AL, Rantanen T, Ripatti S, Rotter JI, Rudan I, Rudnicka AR, Schulz H, Shin SY, Spector TD, Surakka I, Vitart V, Völzke H, Wareham NJ, Warrington NM, Wichmann HE, Wild SH, Wilk JB, Wjst M, Wright AF, Zgaga L, Zemunik T, Pennell CE, Nyberg F, Kuh D, Holloway JW, Boezen HM, Lawlor DA, Morris RW, Probst-Hensch N, Kaprio J, Wilson JF, Hayward C, Kähönen M, Heinrich J, Musk AW, Jarvis DL, Gläser S, Järvelin MR, Ch Stricker BH, Elliott P, O'Connor GT, Strachan DP, London SJ, Hall IP, Gudnason V, Tobin MD. 2011. Genome-Wide Association and Large-Scale Follow Up Identifies 16 New Loci Influencing Lung Function. *Nat Genet* **43**:1082–1090. doi:10.1038/ng.941
- Tabak C, Smit HA, Räsänen L, Fidanza F, Menotti A, Nissinen A, Feskens EJM, Heederik D, Kromhout D. 2001. Alcohol Consumption in Relation to 20-Year COPD Mortality and Pulmonary Function in Middle-Aged Men from Three European Countries. *Epidemiology* **12**:239–245. doi:10.1097/00001648-200103000-00018
- Thompson JR, Minelli C, Del Greco M F. 2016. Mendelian Randomization using Public Data from Genetic Consortia. *Int J Biostat* **12**. doi:10.1515/ijb-2015-0074
- Tillmann T, Vaucher J, Okbay A, Pikhart H, Peasey A, Kubinova R, Pajak A, Tamosiunas A, Malyutina S, Hartwig FP, Fischer K, Veronesi G, Palmer T, Bowden J, Davey Smith G, Bobak M, Holmes M V. 2017. Education and Coronary Heart Disease: Mendelian Randomisation Study. *BMJ* **358**. doi:10.1136/bmj.j3542
- Torvik FA, Eilertsen EM, McAdams TA, Gustavson K, Zachrisson HD, Brandlistuen R, Gjerde LC, Havdahl A, Stoltenberg C, Ask H, Ystrom E. 2020. Mechanisms Linking Parental Educational Attainment with Child ADHD, Depression, and Academic Problems: a Study of Extended Families in The Norwegian Mother, Father and Child Cohort Study. *J Child Psychol Psychiatry Allied Discip*. doi:10.1111/jcpp.13197
- Treur JL, Demontis D, Davey Smith G, Sallis H, Richardson TG, Wiers RW, Børghlum AD, Verweij KJH, Munafò MR. 2019. Investigating Causality between Liability to ADHD and Substance Use, and Liability to Substance Use and ADHD Risk, Using Mendelian Randomization. *Addict Biol* e12849. doi:10.1111/adb.12849

- Wang J, Zhao Q, Bowden J, Hemani G, Davey G, Small DS, Zhang NR. 2020. Causal Inference for Heritable Phenotypic Risk Factors Using Heterogeneous Genetic Instruments. *bioRxiv*.
- Weitzman ER, Chen Y-Y. 2005. The Co-Occurrence of Smoking and Drinking among Young Adults in College: National Survey Results from the United States. *Drug Alcohol Depend* **80**:377–386. doi:10.1016/j.drugalcdep.2005.05.008
- Wilens TE, Adamson J, Sgambati S, Whitley J, Santry A, Monuteaux MC, Biederman J. 2007. Do Individuals with ADHD Self-Medicate with Cigarettes and Substances of Abuse? Results from a Controlled Family Study of ADHD. *Am J Addict* **16**:14–23. doi:10.1080/10550490601082742
- World Health Organization. 2004. ICD-10 : International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, 2nd ed. Geneva: World Health Organization.
- Yavorska OO, Burgess S. 2017. MendelianRandomization: An R package for Performing Mendelian Randomization Analyses Using Summarized Data. *Int J Epidemiol* **46**:1734–1739. doi:10.1093/ije/dyx034
- Zhang L, Reif A, Du Rietz E, Lagerberg T, Butwicki A, D’Onofrio BM, Johnell K, Pedersen NL, Larsson H, Chang Z. 2020. Comedication and Polypharmacy with ADHD Medications in Adults: A Swedish Nationwide Study. *J Atten Disord*. doi:10.1177/1087054720923725
- Zhao Q, Chen Y, Wang J, Small DS. 2019. Powerful Three-Sample Genome-Wide Design and Robust Statistical Inference in Summary-Data Mendelian Randomization. *Int J Epidemiol* **48**:1478–1492. doi:10.1093/ije/dyz142

Supplementary information

Genetic IV selection

Among the 12 independent SNPs for *ADHD* identified in the discovery GWAS, two could not be identified in the replication GWAS. rs7674790 was used as a proxy for rs28411770 ($R^2 = 0.798$), but no valid proxy could be identified for rs7674790. All remaining 11 SNPs were significantly associated with ADHD in the replication GWAS and were in the final genetic IV.

Among the 13 independent SNPs for *alcohol use* identified in the discovery GWAS, two could not be identified in the replication GWAS. rs1629270 was used as a proxy for rs142783062 ($R^2 = 0.988$), but no valid proxy could be identified for rs142997686. Five SNPs were not significantly associated with alcohol use in the replication GWAS, leaving 8 SNPs in the final genetic IV.

Among the 44 independent SNPs for *CAD* identified in the discovery GWAS, ten could not be identified in the replication GWAS. No valid proxies could be identified for chr2:21378433:D, chr2:44074126:D and chr3:138099161:I. rs12118770 was used as a proxy for rs6689306 ($R^2 = 1$), rs17465637 as a proxy for rs67180937 ($R^2 = 1$), rs17740744 as a proxy for rs17678683 ($R^2 = 0.827$), rs6725887 as a proxy for chr2:203828796:I ($R^2 = 0.98$), rs532436 as a proxy for rs2519093 ($R^2 = 0.993$), rs747895 as a proxy for rs7212798 ($R^2 = 1$) and rs5760293 as a proxy for rs180803 ($R^2 = 1$). Five SNPs were not significantly associated with CAD in the replication GWAS, leaving 36 SNPs in the final genetic IV.

Among the 40 independent SNPs for *EA* identified in the discovery GWAS, two could not be identified in the replication GWAS. rs3904683 was used as a proxy for rs71537331 ($R^2 = 0.954$), but no valid proxy could be identified for rs10061788. Two SNPs were not significantly associated with EA in the replication GWAS, leaving 37 SNPs in the final genetic IV.

All the 13 independent SNPs for *COPD* were identified in the discovery GWAS were present in the replication GWAS. Four SNPs were not significantly associated with lung function in the replication GWAS, leaving nine SNPs in the final genetic IV.

All the ten independent SNPs for *smoking* identified in the discovery GWAS were present in the replication GWAS. Five SNPs were not significantly associated with smoking initiation in the replication GWAS; however, since a minimum of six SNPs was required for multivariable MR, one SNP that was not significant in the replication GWAS ($P = 0.065$) was included in the final genetic IV.

Supplementary Table 1: The included genome-wide association studies (GWASs), their samples and sample sizes. ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, SM = Smoking. * = Cases: KORA registry. Controls: KORA S4/F4 study and PopGen. ** = Controls: Berlin Aging Study II (BASE-II). *** = Controls: 1992, 1997, 2002 and 2007 FINRISK surveys.

Exposure	Source	#	Samples	Discovery GWAS		Replication GWAS	
				N (cases / controls)	Total (cases / controls)	Source	N (cases / controls)
ADHD	Demontis et al. (2019)	1	iPSYCH-ADHD, Denmark	14584 / 22492	55374 (20183 / 35191)	Demontis et al. (2019)	55374 (20183 / 35191)
		2	Psychiatric Genomics Consortium (PGC) - Barcelona, Spain	572 / 425			
		3	Psychiatric Genomics Consortium (PGC) - Beijing, China	1012 / 925			
		4	Psychiatric Genomics Consortium (PGC) - Bergen, Norway	295 / 202			
		5	Psychiatric Genomics Consortium (PGC) - Cardiff, UK	721 / 5081			
		6	Psychiatric Genomics Consortium (PGC) - CHOP, USA	262 / 262			
		7	Psychiatric Genomics Consortium (PGC) - Germany	487 / 1290			
		8	Psychiatric Genomics Consortium (PGC) - IMAGE-I	700 / 700			
		9	Psychiatric Genomics Consortium (PGC) - IMAGE-II	624 / 1755			
		10	Psychiatric Genomics Consortium (PGC) - PUWMA	635 / 635			
		11	Psychiatric Genomics Consortium (PGC) - Toronto, Canada	109 / 109			
		12	Psychiatric Genomics Consortium (PGC) - Yale-Penn	182 / 1315			
AU	Kranzler et al. (2019)	1	Million Veteran Program (MVP)	274424	274424	UK Biobank	147267
CAD	Nikpay et al. (2015)	1	Precocious Coronary Artery Disease (PROCARDIS)	5719 / 6545	187599 (61289 / 126310)	UK Biobank	361194 (12769 / 348425)
		2	Helsinki Sudden Death Study/Tampere Coronary Study (HSDS)	206 / 259			
		3	Atherosclerotic Disease, VAscular function, and genetic Epidemiology (ADVANCE)	278 / 312			
		4	Beijing Atherosclerosis Study (BAS)	505 / 1021			
		5	CARDIOGENICS	392 / 410			
		6	China Atherosclerosis Study (CAS)	1010 / 3998			
		7	Cleveland Clinic Gene Bank (CCGB_2)	1628 / 368			
		8	The Corogene Study	2083 / 2048			
		9	Duke Cathgen Study (DUKE_2)	1216 / 653			
		10	Estonian Genome Center of University of Tartu (EGCUT)	658 / 5841			
		11	Functional Genomic Diagnostic Tools for Coronary Artery Disease (FGENTCARD)	1802 / 466			
		12	Genomics Research in Cardiovascular Disease (GENRIC)	2099 / 2690			
		13	German Myocardial Infarction Family Study I (GERMIFS I)	634 / 1608			
		14	German Myocardial Infarction Family Study II (GERMIFS II)	1207 / 1288			
		15	German Myocardial Infarction Family Study III KORA (GERMIFS III KORA) *	1061 / 1467			
		16	German Myocardial Infarction Family Study IV (GERMIFS_IV) **	1089 / 1147			
		17	Genetics of Diabetes and Audit Research in Tayside Scotland (GODARTS)	877 / 2187			

		18	MRC/BHF Heart Protection Study (HPS)	2700 / 2758			
		19	MT. SINAI BioMe Biobank Platform IPM_AA (IPM_AA)	361 / 2778			
		20	MT. SINAI BioMe Biobank Platform IPM_EA (IPM_EA)	487 / 1381			
		21	MT. SINAI BioMe Biobank Platform IPM_HA (IPM_HA)	758 / 3337			
		22	London Life Sciences Prospective Population Study (LOLIPOP)	2791 / 3757			
		23	Ludwigshafen Risk and Cardiovascular Health Study (LURIC)	2095 / 503			
		24	Medstar Cardiac Catheterization Study (MEDSTAR)	933 / 468			
		25	Myocardial Infarction Genetics Consortium (MIGen)	2905 / 2998			
		26	Ottawa Heart Genetic Study A2 (OHGS_A2)	947 / 1008			
		27	Ottawa Heart Genetic Study B2 (OHGS_B2)	1294 / 1529			
		28	Ottawa Heart Genetic Study C2 (OHGS_C2)	843 / 318			
		29	University of Pennsylvania Medical Center Cardiac Catheterization Study (PENNCATH)	933 / 468			
		30	The Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	119 / 830			
		31	A subset of FINRISK cohort study (PREDICTCVD) ***	631 / 334			
		32	Sikh Diabetes Study/ Asian Indian Diabetic Heart Study (SDS/AIDHS)	836 / 761			
		33	The Hellenic Study of Interactions between Snps and Eating in Atherosclerosis Susceptibility (THISEAS)	426 / 594			
		34	TWINGENE	814 / 5999			
		35	The Uppsala Longitudinal Study of Adult Men (ULSAM)	322 / 857			
		36	Wellcome Trust Case Control Consortium (WTCCC)	1926 / 2938			
		37	The Pakistan Risk Of Myocardial Infarction Study 1 (PROMIS1)	4651 / 4452			
		38	The Pakistan Risk Of Myocardial Infarction Study 2 (PROMIS2)	4380 / 3929			
		39	Leipzig Research Center for Civilization Diseases - Heart Study (LIFE-HEART)	1535 / 772			
		40	Women's Genome Health Study (WGHS)	1007 / 22286			
		41	INTERHEART Study (ITH_2)	402 / 448			
		42	Mayo Clinic Vascular Disease Biorepository (MAYO-VDB)	745 / 1389			
		43	Age, Gene/Environment Susceptibility–Reykjavik Study (AGES)	397 / 2474			
		44	The Rotterdam Study (RS)	506 / 5335			
		45	Framingham Heart Study (FHS)	259 / 4202			
		46	Family Heart Study (FamHS)	334 / 3446			
		47	PROspective Study of Pravastatin in the Elderly at Risk (PROSPER)	2034 / 3210			
		48	Atherosclerosis Risk in Communities (ARIC)	454 / 8443			
EA	Lee et al. (2018), excluding British cohorts and 23&Me	1	Age, Gene/ Environment Susceptibility–Reykjavik Study (AGES)	3212	242952	UK Biobank	451247
		2	Austrian Stroke Prevention Study (ASPS)	777			
		3	Berlin Aging Study II (BASE-II)	1619			
		4	Cohorte Lausannoise (CoLaus)	3269			
		5	Copenhagen Studies on Asthma in Childhood 2000 (COPSAC2000)	318			
		6	Croatia Korčula	842			

7	deCODE genetics	46758			
8	Dortmund Health Study (DHS)	953			
9	Estonian Genome Center, University of Tartu (EGCUT)	36631			
10	Erasmus Rucphen Family Study (ERF)	2433			
11	The National FINRISK Study	1685			
12	Finnish Twin Cohort (FTC)	2418			
13	Genetics of Overweight Young Adults (GOYA)	1459			
14	Health 2000 Cases	797			
15	Health 2000 Controls	819			
16	Helsinki Birth Cohort Study (HBCS)	1617			
17	Hunter Community Study (HCS)	1946			
18	Heinz Nixdorf Recall Study (HNRS) CorexB	1401			
19	Heinz Nixdorf Recall Study (HNRS) Oexpr	1347			
20	Heinz Nixdorf Recall Study (HNRS) Omni1	778			
21	Health and Retirement Study (HRS)	9963			
22	Hypergenes	815			
23	Italian Network of Genetic Isolates (INGI) - Carlantino	947			
24	Italian Network of Genetic Isolates (INGI) - Friuli Venezia Giulia	943			
25	Kooperative Gesundheitsforschung in der Region Augsburg (KORA) S3	2655			
26	Kooperative Gesundheitsforschung in der Region Augsburg (KORA) S4	2721			
27	The LifeLines Cohort Study	12539			
28	Minnesota Center for Twin and Family Research (MCTFR)	3819			
29	Molecular Genetics of Schizophrenia (MGS)	2313			
30	Mother and Child Cohort of NIPH (MoBa)	622			
31	Nijmegen Biomedical Study (NBS)	1808			
32	Netherlands Study of Depression and Anxiety (NESDA)	1820			
33	Northern Finland Birth Cohort 1966 (NFBC66)	5297			
34	Netherlands Twin Register (NTR)	5246			
35	Ogliastra Genetic Park (OGP)	370			
36	Ogliastra Genetic Park (OGP) Talana	544			
37	Prevention of Renal and Vascular End-stage Disease (PREVEND)	3578			
38	Queensland Institute of Medical Research (QIMR)	8006			
39	Rotterdam Study Baseline (RS-I)	6108			
40	Rotterdam Study Extension of Baseline (RS-II)	1667			
41	Rotterdam Study Young (RS-III)	3040			
42	Rush University Medical Center - Memory and Aging Project (Rush-MAP)	887			
43	Rush University Medical Center - Religious Orders Study (Rush-ROS)	808			
44	SardiNIA Study of Aging	5616			
45	Study of Health in Pomerania (SHIP)	3556			

		46	Study of Health in Pomerania TREND (SHIP-TREND)	901			
		47	Swedish Twin Registry (STR) - Salty	4832			
		48	Swedish Twin Registry (STR) - Twingene	9553			
		49	The Hellenic Study of Interactions between SNPs & Eating in Atherosclerosis Susceptibility (THISEAS)	829			
		50	The Cardiovascular Risk in Young Finns Study (YFS)	2029			
		51	National Longitudinal Study of Adolescent to Adult Health (Add Health)	5690			
		52	Geisinger Health System	14562			
		53	Wisconsin Longitudinal Study (WLS)	7789			
COPD	Soler Arterias et al. (2011)	1	Age, Gene/Environment Susceptibility (AGES)	1689	94612	UK Biobank	117241
		2	Atherosclerosis Risk in Communities (ARIC)	9078			
		3	British 1958 Birth Cohort–Type 1 Diabetes Genetics Consortium (B58C T1DGC)	2343			
		4	British 1958 Birth Cohort–Wellcome Trust Case Control Consortium (B58C WTCCC)	1372			
		5	Busselton Health Study 1 (BHS1)	1168			
		6	Cardiovascular Health Study (CHS)	3140			
		7	CROATIA-Korcula	825			
		8	CROATIA-Vis	769			
		9	European Community Respiratory Health Survey (ECRHS)	1594			
		10	European Prospective Investigation into Cancer and Nutrition, Obese Cases (EPIC Obsese Cases)	1104			
		11	European Prospective Investigation into Cancer and Nutrition Cohort (EPIC Population Based)	2336			
		12	Framingham Heart Study (FHS)	7911			
		13	Finnish Twin Cohort incorporating FinnTwin16 and FITSA (FTC)	134			
		14	Health, Aging and Body Composition (Health ABC)	1472			
		15	Finnish Health 2000 Survey (Health 2000)	821			
		16	Cooperative Health Research in the Region of Augsburg F4 (KORA F4)	904			
		17	Cooperative Health Research in the Region of Augsburg S3 (KORA S3)	555			
		18	Northern Finland Birth Cohort of 1966 (NFBC1966)	4556			
		19	Orkney Complex Disease Study (ORCADES)	692			
		20	Rotterdam Study I (RS-I)	1224			
		21	Rotterdam Study II (RS-II)	852			
		22	Study of Health in Pomerania (SHIP)	1777			
		23	Twins UK-I	1885			
		24	Adult-Onset Asthma and Nitric Oxide (ADONIX)	1410			
		25	Busselton Health Study 2 (BHS2)	3038			
		26	British Regional Heart Study (BRHS)	3862			
		27	British Women’s Heart and Health Study (BWHHS)	3635			

		28	Coronary Artery Risk Development in Young Adults (CARDIA)	1626			
		29	CROATIA-SPLIT	491			
		30	Gedling	1266			
		31	Generation Scotland: Scottish Family Health Study (GS:SFHS)	10399			
		32	Hertfordshire Cohort Study (HCS)	2848			
		33	Lothian Birth Cohort 1936 (LBC1936)	991			
		34	LifeLines	3078			
		35	Multi-Ethnic Study of Atherosclerosis (MESA-Lung)	1469			
		36	Nottingham Smokers	521			
		37	Medical Research Council National Survey of Health and Development (i.e British 1946 Birth Cohort) (NSHD)	2511			
		38	Rotterdam Study III (RS-III)	1247			
		39	Swiss Study on Air Pollution and Lung Disease in Adults (SAPALDIA)	5646			
SM	Liu et al. (2019), excluding UK Biobank and 23&Me	40	TwinsUK-II	2373	249171	UK Biobank	301733
		1	Avon Longitudinal Study of Parents and Children (ALSPAC)	11345			
		2	The Barrett's and Esophageal Adenocarcinoma. Genetic Susceptibility Study (BEAGESS)	4293			
		3	Brisbane Longitudinal Twin Study (BLS)	1238			
		4	Center on Antisocial Drug Dependence (CADD)	1192			
		5	deCODE Genetics/AMGEN, Inc. (deCODE)	57097			
		6	Estonian Genome Center 370CNV (EGCUT 370CNV)	1758			
		7	Estonian Genome Center Omni (EGCUT Omni)	7041			
		8	Estonian Genome Center Exome (EGCUT Exome)	4418			
		9	Electronic Medical Records and Genomics Cataracts (eMERGE Cataracts)	3873			
		10	Electronic Medical Records and Genomics PAD (eMERGE PAD)	1350			
		11	Framingham Heart Study (FHS)	6177			
		12	Finnish Twin Cohort (FinnTwin)	1006			
		13	Genes for Good (GFG)	1829			
		14	Harvard Affy	7763			
		15	Harvard Illumina	7346			
		16	Health and Retirement Study (HRS)	9989			
		17	The Nord-Trøndelag Health Study (HUNT)	66716			
		18	Minnesota Center for Twin and Family Research (MCTFR)	6181			
		19	Multi-Ethnic Study of Atherosclerosis (MESA)	2165			
		20	Metabolic Syndrome in Men (METSIM)	9607			
		21	Netherlands Study on Cognition, Environment and Genes (NESCOG)	477			
		22	Nicotine Addiction Genetics-Finland study (NAG-FIN)	2052			
		23	Netherlands Twin Register (NTR)	7266			
		24	SardiNIA project (SardiNIA)	5459			

		25	The National Institute of Neurological Disorders and Stroke Genetics Network (STROKE)	3665			
		26	Women's Health Initiative (WHI) MOPMAP	2257			
		27	Women's Health Initiative (WHI) HIPFX	3195			
		28	Women's Health Initiative (WHI) GARNET	4190			
		29	Women's Health Initiative (WHI) LLS	1099			
		30	Women's Health Initiative (WHI) GECCO	1554			
		31	Women's Health Initiative (WHI) WHIMS	5573			

Supplementary Table 2: Cohort overlap between discovery genome-wide association studies (GWASs) of all traits. ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, SM = Smoking.

Trait 1	Trait 2	Overlapping cohorts	Overlap (N)	Overlap (% trait 1)	Overlap (% trait 2)
ADHD	AU	No overlap	0	0	0
ADHD	CAD	No overlap	0	0	0
ADHD	EA	No overlap	0	0	0
ADHD	COPD	No overlap	0	0	0
ADHD	SM	No overlap	0	0	0
AU	CAD	No overlap	0	0	0
AU	EA	No overlap	0	0	0
AU	SM	No overlap	0	0	0
AU	COPD	No overlap	0	0	0
CAD	EA	Age, Gene/Environment Susceptibility (AGES), Estonian Genome Center of University of Tartu (EGCUT), The Hellenic Study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility (THISEAS), Rotterdam Study (EA: I, II, III; CAD: no specification), National FINRISK Study (CAD: controls in PREDICTCVD), Berlin Aging Study II (BASE-II) (CAD: controls in German Myocardial Infarction Family Study (GERMIFS) IV), Cooperative Health Research in the Region of Augsburg (KORA) S4 (CAD: controls in German Myocardial Infarction Study (GERMIFS) III KORA)	18.988	10	8
CAD	COPD	Age, Gene/Environment Susceptibility (AGES), Atherosclerosis Risk in Communities (ARIC), Wellcome Trust Case Control Consortium, Framingham Heart Study (FHS), Rotterdam Study (COPD: I, II; CAD: no specification), Cooperative Health Research in the Region of Augsburg F4 (KORA F4) (CAD: controls in German Myocardial Infarction Study (GERMIFS) III KORA)	17.323	9	18
CAD	SM	Framingham Heart Study (FHS), Estonian Genome Center of University of Tartu (EGCUT) (SM: 370CNV + Omni + Exome; CAD: no specification)	10.960	6	2
EA	COPD	Age, Gene/Environment Susceptibility (AGES), Cooperative Health Research in the Region of Augsburg (KORA) S3, KORA S4 (COPD: F4), CROATIA-Korcula, LifeLines, Northern Finland Birth Cohort 1966, Rotterdam Study I+II+III, Finish Health 2000 Survey (COPD: no specification; EA: cases & controls), Study of Health in Pomerania (SHIP) (COPD: no specification; EA: SHIP + SHIP-TREND)	17.504	7	19
EA	SM	deCODE Genetics, Finnish Twin Cohort (FTC), Health and Retirement Study (HRS), Minnesota Center for Twin and Family Research (MCRFR), Netherlands Twin Register (NTR), Sardinia Study of Aging, Estonian Genome Center University of Tartu (EGCUT) (SM: 370CNV + Omni + Exome, CAD: no specification)	85.468	35	14
COPD	SM	Framingham Heart Study (FHS), MESA (COPD: MESA-Lung, AU: no specification), Finnish Twin Cohort (COPD: incorporating FinnTwin 16 & FITSA, AU: no specification)	7.780	8	1

Supplementary Table 3: The independent single nucleotide polymorphisms (SNPs) for each exposure obtained from the discovery genome-wide association studies (GWASs) and the final genetic instrumental variables (IVs) obtained from the replication GWASs. A1 = effect allele, A2 = other allele, ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, S.E. = standard error, SM = Smoking. x = The value was not reported in the publication. Grey rows indicate SNPs which were not significantly associated with their respective trait ($P > 0.05$) and were excluded from the final genetic IV of that trait. * = This SNP was included in the final genetic IV for smoking in order for the IV to have enough SNPs for multivariable Mendelian Randomization analyses. ** = This proxy was not identified using LDLink, but was referred to in the original publication.

		Discovery GWAS						Replication GWAS					
Exposure		SNP	A1	A2	β	S.E.	P	Proxy?	β	S.E.	P	R ²	
ADHD	1	rs10262192	A	G	0.073	0.013	2.89E-08	-	0.073	0.013	2.89E-08		
	2	rs112984125	A	G	-0.106	0.015	3.58E-13	-	-0.106	0.015	3.58E-13		
	3	rs11591402	A	T	-0.093	0.016	1.34E-08	-	-0.093	0.016	1.34E-08		
	4	rs1222063	A	G	0.096	0.017	3.07E-08	Not identified (<i>R</i> ² > 0.80)					
	5	rs1427829	A	G	0.080	0.013	1.82E-09	-	0.080	0.013	1.82E-09		
	6	rs212178	A	G	-0.115	0.020	7.68E-09	-	-0.115	0.020	7.68E-09		
	7	rs281320	T	G	-0.074	0.013	3.14E-08	-	-0.074	0.013	3.14E-08		
	8	rs28411770	T	C	0.086	0.015	1.15E-08	rs7674790	0.117	0.021	1.39E-08	0.798	
	9	rs4858241	T	G	0.079	0.014	1.74E-08	-	0.079	0.014	1.74E-08		
	10	rs4916723	A	C	-0.077	0.014	1.58E-08	-	-0.077	0.014	1.58E-08		
	11	rs74760947	A	G	-0.180	0.032	1.35E-08	-	-0.180	0.032	1.35E-08		
	12	rs9677504	A	G	0.117	0.021	1.39E-08	-	0.072	0.014	1.80E-07		
AU	1	rs1229984	C	T	x	x	3.62E-133	-	-0.050	0.003	0.000		
	2	rs1260326	C	T	x	x	2.04E-16	-	-0.006	0.001	2.30E-10		
	3	rs12639940	A	G	x	x	3.06E-09	-	0.009	0.001	2.00E-19		
	4	rs13107325	C	T	x	x	2.24E-30	-	-0.013	0.002	1.10E-11		
	5	rs142783062	D	I	x	x	9.50E-23	rs1629270	-0.001	0.001	0.170	0.988	
	6	rs142997686	D	I	x	x	5.39E-09	-	No proxy identified (<i>R</i> > 0.80)				
	7	rs2683616	A	G	x	x	4.95E-10	-	-0.005	0.001	6.10E-07		
	8	rs2961816	A	C	x	x	9.75E-09	-	0.002	0.001	0.038		
	9	rs4423856	T	C	x	x	1.48E-08	-	-0.001	0.001	0.270		
	10	rs4841132	A	G	x	x	3.62E-08	-	-5.01E-04	0.001	0.670		
	11	rs62033408	A	G	x	x	1.11E-19	-	-0.005	0.002	0.004		
	12	rs75723348	T	G	x	x	1.11E-08	-	8.73E-04	0.001	0.390		
	13	rs9902512	C	G	x	x	6.24E-09	-	0.005	0.001	2.90E-05		
CAD	1	chr2:203828796:I	I	D	0.140	0.009	1.45E-60	rs6725887 **	-0.004	4.48E-04	0.000	0.98	
	2	chr2:21378433:D	D	I	0.068	0.012	2.89E-08	Not identified (<i>R</i> ² > 0.80)					
	3	chr2:44074126:D	I	D	0.058	0.010	2.60E-08	Not identified (<i>R</i> ² > 0.80)					
	4	chr3:138099161:I	I	D	0.077	0.013	2.89E-09	Not identified (<i>R</i> ² > 0.80)					
	5	rs10139550	G	C	0.058	0.010	1.38E-08	-	-0.002	4.48E-04	4.17E-05		

6	rs10840293	A	G	0.058	0.010	1.28E-08	-	0.002	4.45E-04	5.98E-04	
7	rs11191416	T	G	0.077	0.013	4.65E-09	-	0.002	7.93E-04	0.022	
8	rs11206510	T	C	0.077	0.014	2.34E-08	-	0.002	5.63E-04	4.50E-04	
9	rs11556924	C	T	0.077	0.012	5.34E-11	-	-0.001	4.51E-04	0.005	
10	rs11838776	A	G	0.068	0.011	1.83E-10	-	0.001	4.96E-04	0.003	
11	rs12202017	A	G	0.068	0.010	1.98E-11	-	0.002	4.50E-04	3.32E-04	
12	rs1412444	T	C	0.068	0.010	5.15E-12	-	0.003	4.85E-04	5.40E-09	
13	rs16986953	A	G	0.086	0.015	1.45E-08	-	0.002	4.65E-04	5.02E-04	
14	rs17087335	T	G	0.058	0.011	4.59E-08	-	0.005	8.80E-04	4.27E-09	
15	rs17678683	G	T	0.095	0.016	3.00E-09	rs17740744	0.005	5.29E-04	0.000	0.8265
16	rs180803	G	T	0.182	0.029	1.64E-10	rs5760293	-0.003	0.001	0.030	1
17	rs1870634	G	T	0.077	0.010	5.55E-15	-	0.001	5.66E-04	0.019	
18	rs2107595	A	G	0.077	0.012	8.05E-11	-	-0.003	4.88E-04	5.97E-10	
19	rs2128739	A	C	0.068	0.010	7.05E-11	-	-1.34E-04	8.27E-04	0.871	
20	rs2487928	A	G	0.058	0.009	4.41E-11	-	-0.002	4.67E-04	1.63E-05	
21	rs2519093	T	C	0.077	0.011	1.19E-11	rs532436	0.003	4.42E-04	1.46E-09	0.9934
22	rs2681472	G	A	0.077	0.012	6.17E-11	-	0.004	6.15E-04	3.15E-11	
23	rs28451064	A	G	0.131	0.016	1.33E-15	-	0.003	4.91E-04	1.12E-10	
24	rs2891168	G	A	0.191	0.009	2.29E-98	-	0.002	4.44E-04	1.30E-06	
25	rs3184504	T	C	0.068	0.011	1.03E-09	-	-0.002	5.86E-04	2.38E-04	
26	rs3918226	T	C	0.131	0.022	1.69E-09	-	0.004	6.64E-04	3.77E-11	
27	rs4252185	C	T	0.293	0.025	1.64E-32	-	-0.008	4.40E-04	0.000	
28	rs4420638	G	A	0.095	0.015	7.07E-11	-	0.002	4.40E-04	5.23E-08	
29	rs4468572	C	T	0.077	0.009	4.44E-16	-	0.004	8.17E-04	2.60E-06	
30	rs4593108	C	G	0.068	0.011	8.82E-10	-	-0.011	7.51E-04	0.000	
31	rs55730499	T	C	0.315	0.024	5.39E-39	-	-0.004	5.60E-04	2.72E-10	
32	rs56062135	C	T	0.068	0.012	4.52E-09	-	-0.003	4.47E-04	3.92E-10	
33	rs56289821	G	A	0.131	0.017	4.44E-15	-	0.002	5.83E-04	2.43E-05	
34	rs56336142	T	C	0.068	0.012	1.85E-08	-	7.86E-05	5.68E-04	0.890	
35	rs663129	A	G	0.058	0.011	3.20E-08	-	0.014	8.06E-04	0.000	
36	rs6689306	A	G	0.058	0.010	2.60E-09	rs12118770	-0.003	5.68E-04	8.62E-07	1
37	rs67180937	G	T	0.077	0.011	1.01E-12	rs17465637	8.39E-04	6.22E-04	0.177	1
38	rs7212798	C	T	0.077	0.014	1.88E-08	rs747895	-0.005	7.61E-04	3.53E-11	1
39	rs72689147	G	T	0.068	0.012	6.07E-09	-	-0.002	5.18E-04	3.48E-06	
40	rs7528419	A	G	0.113	0.011	1.97E-23	-	-0.004	6.81E-04	2.49E-10	
41	rs7568458	A	T	0.058	0.009	3.62E-10	-	2.96E-04	5.41E-04	0.584	
42	rs8042271	G	A	0.095	0.017	3.68E-08	-	0.005	0.002	0.026	
43	rs9349379	G	A	0.131	0.010	1.81E-42	-	7.26E-04	5.20E-04	0.162	
44	rs9970807	C	T	0.122	0.016	5.00E-14	-	-0.004	6.58E-04	5.57E-10	

EA	1	rs10061788	A	G	0.028	0.004	7.81E-12	Not identified ($R^2 > 0.80$)				
	2	rs1035578	A	G	-0.017	0.003	1.19E-08	-	-0.058	0.011	9.18E-08	
	3	rs10761741	T	G	0.018	0.003	1.02E-09	-	0.079	0.011	2.09E-13	
	4	rs10939881	A	G	0.017	0.003	1.18E-08	-	0.097	0.011	1.54E-19	
	5	rs1106761	A	G	-0.017	0.003	3.94E-08	-	-0.087	0.011	6.53E-15	
	6	rs11191193	A	G	0.022	0.003	1.79E-12	-	0.083	0.011	1.42E-13	
	7	rs115000530	A	T	-0.039	0.007	1.53E-08	-	-0.140	0.024	3.21E-09	
	8	rs11581418	T	C	0.027	0.004	9.54E-10	-	0.054	0.016	7.84E-04	
	9	rs11663596	T	C	0.018	0.003	5.48E-10	-	0.051	0.011	3.20E-06	
	10	rs11690172	A	G	0.017	0.003	1.19E-08	-	0.031	0.011	0.004	
	11	rs11811826	A	T	0.022	0.003	2.46E-10	-	0.113	0.013	1.58E-18	
	12	rs11894352	T	G	-0.029	0.005	5.04E-09	-	0.030	0.018	0.088	
	13	rs12643771	T	C	0.021	0.003	8.41E-11	-	0.060	0.012	3.01E-07	
	14	rs12987662	A	C	0.024	0.003	2.02E-15	-	0.117	0.011	9.92E-27	
	15	rs1599381	A	G	0.017	0.003	1.30E-08	-	0.034	0.011	0.002	
	16	rs16826345	T	C	0.044	0.008	3.28E-08	-	0.005	0.028	0.850	
	17	rs1701704	T	G	-0.019	0.003	2.62E-09	-	-0.083	0.011	1.32E-13	
	18	rs17119973	A	G	-0.020	0.003	1.48E-09	-	-0.059	0.012	1.60E-06	
	19	rs1985031	A	G	-0.023	0.003	9.53E-13	-	-0.116	0.012	1.29E-22	
	20	rs2457660	T	C	-0.018	0.003	6.73E-09	-	-0.072	0.011	1.61E-10	
	21	rs2837992	T	C	0.018	0.003	1.03E-09	-	0.063	0.011	1.04E-08	
	22	rs34305371	A	G	0.034	0.005	2.00E-10	-	0.129	0.018	2.03E-13	
	23	rs35121390	C	G	-0.022	0.003	7.11E-13	-	-0.160	0.011	1.95E-48	
	24	rs3890065	C	G	0.017	0.003	4.30E-08	-	0.053	0.011	1.21E-06	
	25	rs4382592	T	G	-0.017	0.003	4.37E-08	-	0.087	0.011	5.00E-15	
	26	rs4678311	T	C	0.017	0.003	1.62E-08	-	-0.074	0.012	2.06E-10	
	27	rs55830725	A	T	-0.021	0.004	4.86E-08	-	0.034	0.011	0.002	
	28	rs61358955	A	T	0.020	0.003	2.80E-10	-	-0.041	0.014	0.005	
	29	rs62103084	T	C	0.029	0.005	1.07E-08	-	0.053	0.011	4.05E-06	
	30	rs6430030	T	C	0.018	0.003	2.26E-09	-	0.062	0.017	4.18E-04	
	31	rs71537331	T	C	-0.018	0.003	3.02E-08	rs3904683	-0.053	0.011	7.38E-07	0.9542
	32	rs7306755	A	G	0.023	0.004	1.97E-10	-	0.073	0.011	1.21E-11	
	33	rs7565794	T	C	-0.017	0.003	1.49E-08	-	0.108	0.013	4.00E-16	
	34	rs7613360	T	C	-0.027	0.003	5.94E-19	-	-0.067	0.012	9.95E-09	
	35	rs77609760	A	G	0.032	0.006	2.92E-08	-	-0.145	0.011	2.20E-40	
	36	rs7830404	A	T	-0.017	0.003	2.28E-08	-	0.156	0.021	2.19E-13	
	37	rs7849487	T	G	-0.020	0.003	1.44E-10	-	-0.026	0.011	0.016	
	38	rs7868984	T	C	-0.024	0.003	5.19E-15	-	-0.072	0.011	1.49E-10	
	39	rs895606	A	G	0.018	0.003	1.09E-09	-	-0.147	0.011	7.63E-42	

	40	rs972061	C	G	-0.018	0.003	3.02E-10	-	0.061	0.011	1.28E-08	
COPD	1	rs1036429	T	x	0.038	0.006	2.30E-11	-	0.010	0.005	0.054	
	2	rs11172113	T	x	-0.032	0.006	1.24E-08	-	-0.007	0.004	0.090	
	3	rs12447804	T	x	-0.038	0.007	3.59E-08	-	-0.011	0.005	0.025	
	4	rs12477314	T	x	0.041	0.006	1.68E-12	-	0.027	0.005	1.29E-07	
	5	rs1529672	C	x	-0.048	0.006	3.97E-14	-	0.013	0.005	0.018	
	6	rs153916	T	x	-0.031	0.005	2.12E-08	-	4.06E-05	0.004	0.992	
	7	rs2284746	G	x	-0.040	0.005	7.50E-16	-	-0.009	0.004	0.023	
	8	rs2798641	T	x	-0.041	0.007	8.35E-09	-	-0.033	0.005	7.59E-10	
	9	rs2857595	G	x	0.037	0.006	2.28E-10	-	-0.029	0.005	1.16E-08	
	10	rs2865531	T	x	0.031	0.005	1.77E-11	-	-0.025	0.004	3.79E-09	
	11	rs7068966	T	x	0.033	0.005	6.13E-13	-	0.030	0.004	3.31E-13	
	12	rs993925	T	x	0.034	0.006	1.16E-08	-	0.009	0.004	0.045	
	13	rs9978142	T	x	-0.043	0.008	2.65E-08	-	0.004	0.006	0.462	
SM	1	rs12441907	C	A	-0.046	0.007	1.90E-10	-	-0.002	0.001	0.197	
	2	rs1355334	T	A	0.034	0.006	1.36E-08	-	0.002	0.001	0.064	
	3	rs1979004	T	A	0.035	0.006	8.66E-09	-	-4.30E-04	9.91E-04	0.664	
	4	rs2162965	C	T	0.037	0.007	2.76E-08	-	0.003	0.001	0.016	
	5	rs2186874	T	C	0.042	0.006	2.99E-13	-	-0.005	9.70E-04	2.84E-08	
	6	rs3001723 *	G	A	0.038	0.006	1.07E-09	-	0.003	0.001	0.007	
	7	rs6756212	C	T	-0.035	0.006	5.03E-10	-	-0.005	9.45E-04	7.15E-07	
	8	rs7613360	C	T	0.034	0.006	5.33E-09	-	3.75E-04	9.62E-04	0.697	
	9	rs77311064	G	A	0.055	0.010	3.18E-08	-	-0.001	0.002	0.510	
	10	rs883323	C	T	-0.033	0.006	2.56E-08	-	-0.002	9.80E-04	0.030	

Supplementary Table 4: The summary statistics of the genetic instrumental variables (IVs) for each exposure in the replication genome-wide association studies (GWASs). Grey rows indicate single nucleotide polymorphisms (SNPs) which were not significantly associated with their respective trait ($P > 0.05$) and were excluded from the final genetic IV of that trait. * This SNP was included in the final genetic IV for smoking in order for the IV to have enough SNPs for multivariable Mendelian Randomization analyses. A1 = effect allele, A2 = other allele, ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, S.E. = standard error, SM = Smoking.

Exposure	SNP	CHR	BP	A1	A2	ADHD			AU			CAD			EA			COPD			SM		
						β	S.E.	P	β	S.E.	P	β	S.E.	P	β	S.E.	P	β	S.E.	P	β	S.E.	P
ADHD	rs10262192	7	114091753	A	G				-0.002	9.99E-04	0.088	3.40E-04	4.44E-04	0.443	-0.022	0.011	0.043	-0.004	0.004	0.398	-7.45E-04	9.55E-04	0.435
	rs112984125	1	44173423	A	G				0.002	0.001	0.094	-5.11E-04	4.81E-04	0.288	0.100	0.012	0.000	0.004	0.004	0.422	-0.002	0.001	0.041
	rs11591402	10	106747354	A	T				0.002	0.001	0.058	-8.64E-04	5.43E-04	0.112	0.026	0.013	0.049	0.006	0.005	0.243	-0.002	0.001	0.157
	rs1427829	12	89760744	A	G				7.74E-04	9.94E-04	0.440	9.52E-04	4.42E-04	0.031	-0.050	0.011	2.92E-06	-0.006	0.004	0.164	7.56E-04	9.45E-04	0.424
	rs212178	16	72578131	A	G				-4.85E-04	0.002	0.760	-9.21E-04	7.00E-04	0.188	0.051	0.017	0.003	-0.003	0.007	0.688	-0.004	0.002	0.005
	rs281320	15	47769424	T	G				-0.002	9.92E-04	0.061	5.52E-04	4.40E-04	0.210	0.061	0.011	1.13E-08	-0.007	0.004	0.081	-0.003	9.42E-04	0.001
	rs4858241	3	20669071	T	G				-2.20E-04	0.001	0.830	6.80E-04	4.56E-04	0.136	-0.022	0.011	0.051	-0.003	0.004	0.522	-5.28E-04	9.78E-04	0.589
	rs4916723	5	87854395	A	C				0.005	0.001	7.80E-06	0.001	4.50E-04	0.022	0.061	0.011	3.25E-08	2.45E-04	0.004	0.954	-0.002	9.73E-04	0.035
	rs74760947	8	34352610	A	G				0.002	0.002	0.530	3.16E-04	0.001	0.759	-0.020	0.025	0.434	0.022	0.010	0.023	-0.005	0.002	0.022
	rs9677504	2	215181889	A	G				-7.42E-04	0.001	0.470	-6.42E-04	4.56E-04	0.159	0.002	0.011	0.863	4.21E-04	0.004	0.921	0.002	9.74E-04	0.022
AU	rs7674790	4	31149277	A	T				-0.001	0.002	0.560	-3.00E-04	7.61E-04	0.694	0.019	0.019	0.296	-0.012	0.007	0.083	7.17E-04	0.002	0.661
	rs1229984	4	100239319	C	T	0.019	0.038	0.623				-0.004	0.001	0.008	0.113	0.034	8.84E-04	0.034	0.014	0.016	0.003	0.003	0.341
	rs1260326	2	27730940	C	T	-0.014	0.014	0.289				0.001	4.50E-04	0.008	0.024	0.011	0.025	0.006	0.004	0.174	-4.76E-04	9.61E-04	0.620
	rs12639940	4	39420981	G	A	0.019	0.014	0.166				9.60E-06	4.58E-04	0.983	-0.028	0.011	0.010	-0.006	0.004	0.160	-6.91E-04	9.80E-04	0.481
	rs13107325	4	103188709	T	C	0.020	0.031	0.514				-9.71E-04	8.38E-04	0.246	-0.118	0.020	5.97E-09	0.029	0.008	2.20E-04	-0.002	0.002	0.311
	rs2683616	2	58035555	A	G	-0.007	0.014	0.612				-2.53E-04	4.46E-04	0.571	0.026	0.011	0.017	0.012	0.004	0.004	-0.001	9.51E-04	0.160
	rs2961816	5	50443691	A	C	0.021	0.014	0.130				2.02E-04	4.57E-04	0.659	0.008	0.011	0.445	0.001	0.004	0.739	2.15E-04	9.73E-04	0.825
	rs4423856	4	150984857	C	T	0.008	0.014	0.598				1.21E-04	4.84E-04	0.802	-0.029	0.012	0.015	-0.003	0.005	0.487	0.002	0.001	0.033
	rs4841132	8	9183596	G	A	-0.064	0.016	8.52E-05				-0.001	5.29E-04	0.042	-0.006	0.013	0.627	-0.009	0.005	0.080	0.001	0.001	0.354
	rs62033408	16	53827962	G	A	0.025	0.023	0.285				-0.001	7.64E-04	0.091	0.028	0.019	0.137	0.003	0.007	0.654	-0.001	0.002	0.408
CAD	rs75723348	22	41420679	G	T	-0.011	0.014	0.439				-1.10E-04	4.51E-04	0.807	-0.012	0.011	0.282	0.006	0.004	0.184	-0.003	9.63E-04	0.008
	rs9902512	17	47094274	G	C	0.004	0.015	0.782				-6.41E-04	4.87E-04	0.189	-0.016	0.012	0.182	-7.78E-04	0.005	0.864	-0.001	0.001	0.241
	rs1629270	4	100271306	G	A	0.006	0.014	0.683				0.002	4.60E-04	7.37E-04	-0.042	0.011	1.88E-04	6.35E-05	0.004	0.988	2.57E-05	9.82E-04	0.979
	rs10139550	14	100145710	G	C	-0.041	0.014	0.003	9.82E-04	0.001	0.330				-0.009	0.011	0.414	-0.004	0.004	0.374	0.001	9.65E-04	0.290
	rs10840293	11	9751196	A	G	-0.004	0.014	0.762	9.85E-04	0.001	0.320				-0.010	0.011	0.367	0.002	0.004	0.681	-1.33E-04	9.55E-04	0.889
	rs11191416	10	104604916	G	T	0.053	0.022	0.014	-0.002	0.002	0.340				0.003	0.019	0.868	0.017	0.007	0.023	6.23E-04	0.002	0.713
	rs11206510	1	55496039	C	T	0.010	0.017	0.547	2.08E-04	0.001	0.870				0.016	0.014	0.254	5.95E-04	0.005	0.910	-4.40E-04	0.001	0.714
	rs11556924	7	129663496	T	C	0.006	0.015	0.693	-0.002	0.001	0.140				-0.018	0.011	0.096	0.009	0.004	0.043	-0.004	9.63E-04	1.84E-05
	rs11838776	13	111040681	A	G	0.020	0.015	0.194	0.002	0.001	0.051				0.025	0.012	0.040	-0.005	0.005	0.254	-7.41E-04	0.001	0.489
	rs12202017	6	134173151	G	A	0.006	0.013	0.681	-0.002	0.001	0.120				-0.010	0.011	0.358	0.007	0.004	0.100	-3.59E-04	9.62E-04	0.709
	rs1412444	10	91002927	T	C	-0.007	0.015	0.611	4.65E-04	0.001	0.670				-0.012	0.012	0.311	-0.003	0.005	0.463	6.71E-04	0.001	0.518
COPD	rs16986953	2	19942473	A	G	0.010	0.014	0.498	-4.14E-04	0.001	0.690				5.07E-04	0.011	0.964	0.005	0.004	0.262	7.25E-04	9.92E-04	0.465
	rs17087335	4	57838583	T	G	-0.020	0.024	0.400	0.001	0.002	0.450				4.54E-04	0.021	0.983	0.014	0.008	0.083	5.47E-04	0.002	0.772
	rs1870634	10	44480811	G	T	0.019	0.016	0.240	6.47E-04	0.001	0.610				-0.006	0.014	0.642	0.004	0.005	0.414	0.002	0.001	0.055
	rs2107595	7	19049388	A	G	1.00E-04	0.016	0.997	-2.39E-05	0.001	0.980				0.005	0.012	0.656	0.007	0.005	0.136	-0.002	0.001	0.125
	rs2128739	11	103673277	C	A	-0.026	0.026	0.304	7.42E-04	0.002	0.690				0.023	0.020	0.244	-5.90E-04	0.008	0.939	0.002	0.002	0.200
	rs2487928	10	30323892	A	G	0.004	0.014	0.787	-0.001	0.001	0.210				-7.34E-04	0.011	0.948	-0.001	0.004	0.787	2.80E-04	9.96E-04	0.778
	rs2681472	12	90008959	G	A	0.025	0.018	0.156	-5.97E-04	0.001	0.660				-0.050	0.015	8.30E-04	0.004	0.006	0.507	-5.77E-04	0.001	0.661
	rs28451064	21	35593827	A	G	0.017	0.015	0.259	0.002	0.001	0.058				-0.002	0.012	0.841	0.003	0.005	0.473	6.88E-04	0.001	0.511
	rs2891168	9	22098619	G	A	0.012	0.013	0.378	0.002	0.001	0.036				0.016	0.011	0.150	-0.008	0.004	0.048	0.002	9.55E-04	0.020

	rs3184504	12	111884608	C	T	-0.036	0.018	0.043	-0.001	0.001	0.380				0.029	0.014	0.039	-0.002	0.005	0.653	-2.57E-04	0.001	0.837
	rs3918226	7	150690176	T	C	0.008	0.023	0.739	0.003	0.001	0.047				0.007	0.016	0.655	-0.011	0.006	0.064	0.002	0.001	0.164
	rs4252185	6	161123451	C	T	0.006	0.013	0.660	-6.54E-05	9.89E-04	0.950				-0.005	0.011	0.650	5.51E-04	0.004	0.893	-2.07E-05	9.39E-04	0.982
	rs4420638	19	45422946	G	A	-0.013	0.014	0.341	3.47E-04	9.90E-04	0.730				-0.005	0.011	0.671	0.001	0.004	0.751	0.002	9.39E-04	0.072
	rs4468572	15	79124475	C	T	-0.008	0.027	0.762	-0.002	0.002	0.180				-0.048	0.020	0.018	-0.004	0.008	0.600	-8.68E-04	0.002	0.624
	rs4593108	4	148281001	G	C	0.004	0.025	0.872	-8.88E-04	0.002	0.600				0.014	0.019	0.450	0.002	0.007	0.741	-0.002	0.002	0.148
	rs55730499	6	161005610	T	C	-0.012	0.019	0.522	0.004	0.001	0.005				-0.017	0.014	0.199	-0.016	0.005	0.002	-0.001	0.001	0.349
	rs56062135	15	67455630	T	C	0.019	0.014	0.168	-0.001	0.001	0.200				-0.011	0.011	0.311	-0.011	0.004	0.006	-0.005	9.58E-04	1.45E-07
	rs56289821	19	11188247	A	G	-0.012	0.017	0.492	0.001	0.001	0.310				0.025	0.014	0.078	-8.85E-04	0.005	0.871	0.001	0.001	0.295
	rs56336142	6	39134099	C	T	0.021	0.016	0.197	5.67E-04	0.001	0.650				-0.026	0.014	0.058	-0.023	0.005	1.73E-05	-3.80E-04	0.001	0.753
	rs663129	18	57838401	A	G	-0.002	0.026	0.933	0.004	0.002	0.038				-0.004	0.020	0.833	3.21E-04	0.008	0.966	0.001	0.002	0.426
	rs72689147	4	156639888	T	G	0.004	0.016	0.796	-0.001	0.001	0.280				-0.006	0.013	0.648	-0.004	0.005	0.373	-3.14E-04	0.001	0.776
	rs7528419	1	109817192	G	A	0.007	0.022	0.747	-5.99E-04	0.002	0.700				-0.034	0.017	0.038	0.003	0.006	0.590	0.001	0.001	0.343
	rs7568458	2	85788175	A	T	0.013	0.016	0.437	-4.47E-04	0.001	0.710				-0.021	0.013	0.119	0.008	0.005	0.105	-0.002	0.001	0.111
	rs8042271	15	89574218	A	G	0.031	0.050	0.537	-0.004	0.005	0.370				0.015	0.051	0.765	3.86E-04	0.020	0.984	0.001	0.004	0.746
	rs9349379	6	12903957	G	A	0.007	0.015	0.665	-4.86E-05	0.001	0.970				0.024	0.013	0.054	-0.025	0.005	2.80E-07	-0.002	0.001	0.140
	rs9970807	1	56965664	T	C	0.065	0.020	0.001	-0.001	0.001	0.460				-0.011	0.016	0.494	0.002	0.006	0.722	-3.24E-05	0.001	0.982
	rs12118770	1	154397589	C	T	-0.010	0.017	0.546	0.001	0.001	0.250				0.026	0.014	0.054	0.014	0.005	0.007	-0.002	0.001	0.093
	rs17465637	1	222823529	C	A	-0.032	0.018	0.078	0.001	0.001	0.290				0.015	0.015	0.332	0.004	0.006	0.477	-5.89E-04	0.001	0.657
	rs17740744	2	145295362	G	A	-0.032	0.016	0.041	-0.002	0.001	0.090				-0.006	0.013	0.648	0.007	0.005	0.131	-0.002	0.001	0.141
	rs532436	9	136149830	A	G	0.031	0.013	0.020	4.42E-04	9.96E-04	0.660				-0.001	0.011	0.895	-0.006	0.004	0.121	6.62E-04	9.47E-04	0.484
	rs5760293	22	24662171	G	T	0.011	0.032	0.731	0.002	0.003	0.520				-0.022	0.029	0.444	-0.002	0.011	0.869	-0.002	0.003	0.457
	rs6725887	2	203745885	C	T	-0.006	0.014	0.692	-0.004	0.001	2.50E-04				-0.017	0.011	0.114	-0.009	0.004	0.037	-0.001	9.58E-04	0.232
	rs747895	17	59011296	A	T	-0.001	0.024	0.960	0.002	0.002	0.210				-0.048	0.018	0.009	-0.013	0.007	0.075	-0.002	0.002	0.294
EA	rs1035578	16	12531365	G	A	0.017	0.013	0.216	0.003	9.98E-04	0.011	7.15E-04	4.44E-04	0.107				0.004	0.004	0.287	2.76E-04	9.51E-04	0.771
	rs10761741	10	65066186	G	T	-0.017	0.013	0.217	0.002	0.001	0.053	5.02E-04	4.47E-04	0.261				-0.011	0.004	0.009	3.14E-05	9.53E-04	0.974
	rs10939881	5	60382973	G	A	-0.001	0.013	0.916	2.31E-04	9.95E-04	0.820	-7.98E-05	4.43E-04	0.857				-0.002	0.004	0.570	-0.002	9.48E-04	0.059
	rs1106761	8	142619234	G	A	0.004	0.014	0.767	-0.004	0.001	8.70E-05	-9.16E-04	4.58E-04	0.045				-0.005	0.004	0.275	-0.001	9.88E-04	0.166
	rs11191193	10	103802408	A	G	-0.033	0.014	0.018	1.07E-04	0.001	0.920	-3.57E-04	4.62E-04	0.439				-0.006	0.004	0.170	-9.45E-04	9.92E-04	0.341
	rs115000530	3	69930625	A	T	-0.018	0.032	0.558	-0.003	0.002	0.130	0.001	9.69E-04	0.211				2.50E-04	0.009	0.978	5.30E-04	0.002	0.799
	rs11581418	1	8419766	C	T	0.005	0.023	0.831	-0.002	0.001	0.120	-4.81E-04	6.57E-04	0.464				0.008	0.006	0.199	-8.72E-04	0.001	0.538
	rs11663596	18	50507006	T	C	-0.017	0.014	0.218	0.001	0.001	0.250	-4.29E-04	4.55E-04	0.345				-0.006	0.004	0.129	0.002	9.72E-04	0.062
	rs11690172	2	57387094	A	G	3.00E-04	0.014	0.980	-0.002	0.001	0.088	5.06E-04	4.51E-04	0.261				0.006	0.004	0.137	1.14E-07	9.59E-04	1.000
	rs11811826	1	204603861	T	A	-0.031	0.016	0.054	-0.003	0.001	0.024	3.27E-04	5.29E-04	0.536				0.014	0.005	0.004	8.71E-04	0.001	0.441
	rs11894352	2	47665619	T	G	-0.025	0.021	0.226	-5.02E-04	0.002	0.760	4.24E-04	7.34E-04	0.564				3.21E-04	0.007	0.963	-5.34E-04	0.002	0.734
	rs12643771	4	140753103	C	T	-0.047	0.016	0.002	0.003	0.001	0.018	2.76E-04	4.77E-04	0.563				0.006	0.004	0.192	-0.003	0.001	7.88E-04
	rs12987662	2	100821548	C	A	-0.024	0.013	0.077	-9.91E-04	0.001	0.330	-8.86E-04	4.49E-04	0.049				0.004	0.004	0.372	5.90E-04	9.61E-04	0.539
	rs1599381	3	160867680	G	A	-0.018	0.013	0.180	0.001	9.93E-04	0.140	1.15E-04	4.42E-04	0.795				-0.004	0.004	0.304	0.002	9.48E-04	0.013
	rs16826345	2	185818286	T	C	-0.049	0.039	0.206	0.003	0.003	0.330	0.002	0.001	0.054				0.009	0.011	0.405	0.001	0.002	0.621
	rs1701704	12	56412487	T	G	0.023	0.014	0.102	-8.78E-04	0.001	0.400	-5.56E-04	4.63E-04	0.230				0.016	0.004	1.81E-04	0.005	9.88E-04	2.21E-06
	rs17119973	14	84913111	G	A	0.023	0.015	0.117	7.22E-04	0.001	0.530	5.16E-04	5.04E-04	0.306				0.004	0.005	0.356	-4.01E-05	0.001	0.970
	rs1985031	13	58391264	G	A	0.048	0.015	0.001	-8.46E-04	0.001	0.450	5.26E-04	4.91E-04	0.284				-0.008	0.005	0.096	8.12E-05	0.001	0.938
	rs2457660	2	60757419	C	T	0.012	0.014	0.400	-0.003	0.001	0.007	7.07E-05	4.61E-04	0.878				-6.73E-04	0.004	0.876	-3.82E-04	9.86E-04	0.699
	rs2837992	21	42620520	T	C	-0.018	0.014	0.188	8.62E-04	0.001	0.400	-9.79E-05	4.57E-04	0.830				0.009	0.004	0.039	-9.79E-04	9.75E-04	0.316
	rs34305371	1	72733610	G	A	-0.058	0.026	0.023	-0.004	0.002	0.009	1.08E-04	7.27E-04	0.882				0.002	0.007	0.753	-0.003	0.002	0.083
	rs35121390	6	98399208	C	G	0.017	0.014	0.208	-0.004	0.001	4.10E-04	3.18E-04	4.51E-04	0.480				-0.004	0.004	0.293	6.29E-04	9.65E-04	0.515
	rs3890065	11	12758767	C	G	-0.018	0.014	0.186	2.18E-04	0.001	0.830	-5.16E-04	4.48E-04	0.249				0.017	0.004	5.05E-05	-7.47E-04	9.63E-04	0.438
	rs4382592	9	134870755	T	G	-0.024	0.014	0.081	7.29E-04	0.001	0.480	4.37E-04	4.61E-04	0.343				-0.012	0.004	0.005	0.002	9.83E-04	0.033
	rs4678311	3	137075329	C	T	0.043	0.014	0.002	7.92E-04	0.001	0.460	2.54E-04	4.80E-04	0.598				-0.006	0.004	0.162	6.38E-04	0.001	0.535
	rs55830725	2	237056854	T	A	-0.006	0.013	0.662	4.65E-04	9.91E-04	0.640	-1.86E-06	4.40E-04	0.997				0.005	0.004	0.273	0.002	9.41E-04	0.101

	rs61358955	5	120093858	A	T	-0.002	0.018	0.894	0.001	0.001	0.340	-4.98E-04	5.93E-04	0.401				-0.009	0.006	0.124	-8.15E-05	0.001	0.949
	rs62103084	18	25637562	C	T	-0.011	0.014	0.461	-0.001	0.001	0.220	4.86E-06	4.75E-04	0.992				-0.004	0.004	0.321	8.28E-04	0.001	0.414
	rs6430030	2	144166523	C	T	-0.001	0.023	0.964	-0.001	0.002	0.410	-6.23E-04	7.15E-04	0.384				-0.003	0.007	0.678	7.87E-04	0.002	0.610
	rs7306755	12	123767929	G	A	-0.014	0.014	0.312	-0.002	0.001	0.037	-2.42E-04	4.46E-04	0.587				0.003	0.004	0.516	-0.003	9.52E-04	0.007
	rs7565794	2	162859000	T	C	-0.031	0.017	0.069	0.005	0.001	5.80E-05	-5.31E-04	5.48E-04	0.332				-0.010	0.005	0.064	-7.82E-04	0.001	0.504
	rs7613360	3	49916710	C	T	0.002	0.014	0.866	8.82E-04	0.001	0.410	-5.17E-04	4.78E-04	0.280				0.004	0.004	0.376	0.002	0.001	0.074
	rs77609760	2	142358950	A	G	0.057	0.014	3.81E-05	-0.001	0.001	0.270	6.38E-04	4.50E-04	0.156				-0.016	0.004	1.66E-04	3.75E-04	9.62E-04	0.697
	rs7830404	8	19357346	A	T	-0.042	0.026	0.105	0.002	0.002	0.230	-7.99E-04	8.79E-04	0.363				0.008	0.008	0.319	-0.001	0.002	0.433
	rs7849487	9	1757274	T	G	-2.00E-04	0.014	0.988	-0.002	0.001	0.078	-4.46E-04	4.52E-04	0.324				0.005	0.004	0.242	4.11E-04	9.67E-04	0.671
	rs7868984	9	23357826	T	C	-0.004	0.014	0.787	0.002	0.001	0.120	0.001	4.65E-04	0.003				0.003	0.004	0.547	-0.002	9.93E-04	0.058
	rs895606	9	88003668	A	G	0.007	0.014	0.611	3.11E-04	0.001	0.760	8.93E-05	4.46E-04	0.841				-0.010	0.004	0.021	-0.002	9.54E-04	0.013
	rs972061	5	57545308	C	G	-0.016	0.013	0.239	0.002	9.94E-04	0.056	2.05E-04	4.42E-04	0.643				0.002	0.004	0.628	0.002	9.43E-04	0.009
	rs3904683	1	243416525	T	C	0.013	0.013	0.318	7.68E-04	9.92E-04	0.440	-4.44E-04	4.43E-04	0.316				-0.001	0.004	0.723	-5.99E-04	9.42E-04	0.525
COPD	rs1036429	12	96271428	C	T	0.026	0.016	0.117	-0.002	0.001	0.210	-2.91E-04	5.41E-04	0.591	8.23E-04	0.013	0.950				-3.09E-04	0.001	0.789
	rs11172113	12	57527283	C	T	-8.00E-04	0.014	0.953	-0.002	0.001	0.014	-0.001	4.47E-04	0.023	0.017	0.011	0.121				3.25E-04	9.54E-04	0.733
	rs12447804	16	58075282	T	C	-0.023	0.016	0.151	2.11E-04	0.001	0.860	-4.73E-04	5.30E-04	0.372	-0.003	0.013	0.796				0.002	0.001	0.040
	rs12477314	2	239877148	T	C	0.006	0.017	0.714	6.68E-04	0.001	0.590	3.77E-04	5.53E-04	0.495	0.013	0.013	0.335				-7.98E-04	0.001	0.498
	rs1529672	3	25520582	A	C	-0.006	0.018	0.758	-0.002	0.001	0.180	1.87E-04	5.83E-04	0.748	0.028	0.014	0.050				5.29E-04	0.001	0.672
	rs153916	5	95036700	T	C	-0.001	0.013	0.923	0.001	9.96E-04	0.250	-4.38E-05	4.43E-04	0.921	-0.010	0.011	0.364				2.00E-04	9.45E-04	0.832
	rs2284746	1	17306675	G	C	-0.031	0.013	0.020	-0.002	9.91E-04	0.130	6.95E-04	4.40E-04	0.114	-0.004	0.011	0.674				-5.51E-04	9.40E-04	0.558
	rs2798641	6	109268050	T	C	4.00E-04	0.018	0.984	9.49E-04	0.001	0.460	-8.20E-06	5.69E-04	0.989	0.010	0.014	0.482				-0.003	0.001	0.006
	rs2857595	6	31568469	A	G	0.008	0.016	0.613	-1.39E-05	0.001	0.990	1.65E-04	5.37E-04	0.758	-0.004	0.013	0.749				1.73E-04	0.001	0.880
	rs2865531	16	75390316	A	T	-0.015	0.013	0.259	5.91E-04	0.001	0.560	0.002	4.48E-04	5.14E-07	-0.005	0.011	0.646				4.88E-05	9.58E-04	0.959
	rs7068966	10	12277992	T	C	-0.014	0.013	0.287	-0.001	9.89E-04	0.160	-0.002	4.40E-04	1.63E-04	-0.018	0.011	0.088				0.001	9.41E-04	0.216
	rs993925	1	218860068	T	C	0.014	0.014	0.318	9.13E-04	0.001	0.380	-6.33E-04	4.62E-04	0.171	-0.008	0.011	0.491				-3.23E-04	9.84E-04	0.742
	rs9978142	21	35652239	T	A	-0.006	0.019	0.750	-0.004	0.001	0.003	-0.002	6.22E-04	3.18E-04	-0.007	0.015	0.621				-0.001	0.001	0.275
SM	rs12441907	15	83922387	C	A	0.013	0.017	0.426	0.001	0.001	0.340	9.05E-04	5.67E-04	0.110	0.043	0.014	0.002	0.016	0.005	0.002			
	rs1355334	4	140887493	T	A	0.045	0.014	0.001	-0.003	0.001	0.007	0.001	4.70E-04	0.030	-0.051	0.011	7.79E-06	-0.020	0.004	8.02E-06			
	rs1979004	2	155713421	T	A	0.006	0.014	0.647	3.44E-04	0.001	0.740	1.67E-04	4.63E-04	0.719	-7.16E-04	0.011	0.949	0.005	0.004	0.213			
	rs2162965	5	154797752	C	T	0.049	0.015	0.001	-9.56E-05	0.001	0.930	3.69E-04	5.09E-04	0.469	-0.041	0.012	7.71E-04	-0.004	0.005	0.437			
	rs2186874	11	112912724	T	C	-0.035	0.014	0.011	-0.002	0.001	0.057	-2.22E-04	4.52E-04	0.623	0.056	0.011	3.43E-07	0.017	0.004	4.91E-05			
	rs3001723 *	1	44037685	G	A	0.089	0.014	3.63E-10	-4.51E-04	0.001	0.680	-2.94E-05	4.80E-04	0.951	-0.116	0.012	0.000	-0.010	0.004	0.031			
	rs6756212	2	146140132	C	T	-0.039	0.013	0.003	-2.58E-04	9.94E-04	0.800	1.13E-05	4.42E-04	0.980	-0.007	0.011	0.497	0.017	0.004	4.08E-05			
	rs7613360	3	49916710	C	T	0.057	0.014	3.81E-05	-0.001	0.001	0.270	6.38E-04	4.50E-04	0.156	-0.145	0.011	0.000	-0.016	0.004	1.66E-04			
	rs77311064	4	68020214	G	A	0.011	0.022	0.609	0.002	0.002	0.290	6.85E-04	7.42E-04	0.356	-0.090	0.018	4.99E-07	0.005	0.007	0.466			
	rs883323	5	166987744	C	T	-0.038	0.014	0.006	-0.002	0.001	0.062	-0.001	4.56E-04	0.009	-0.010	0.011	0.367	0.012	0.004	0.005			

Supplementary Table 5: Robustness measures of univariable Mendelian Randomization. ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, Cochran's Q_{EGGER} = Cochran's Q statistic of the MR-Egger regression, Cochran's Q_{IVW} = Cochran's Q statistic of the inverse-variance weighted method, S.E. = standard error, SM = Smoking. * = $P < 0.05$.

Exposure	Outcome	Cochran's Q				MR-Egger Intercept		
		Cochran's Q_{IVW}	P	Cochran's Q_{EGGER}	P	Intercept estimate	S.E.	P
ADHD	AU	27.239	0.002 *	27.203	0.001 *	-2.94E-04	0.003	0.912
ADHD	CAD	20.97	0.021 *	20.37	0.016 *	-5.25E-04	0.001	0.607
ADHD	EA	79.016	7.83E-13 *	72.295	5.39E-12 *	-0.043	0.047	0.36
ADHD	COPD	12.258	0.268	8.565	0.478	0.012	0.006	0.055
ADHD	SM	17.249	0.069	16.432	0.058	-0.001	0.002	0.503
AU	ADHD	7.238	0.405	4.432	0.618	0.013	0.008	0.094
AU	CAD	17.784	0.013 *	14.026	0.029 *	-5.40E-04	4.26E-04	0.205
AU	EA	56.37	7.97E-10 *	55.981	2.94E-10 *	-0.004	0.02	0.838
AU	COPD	10.093	0.183	10.014	0.124	-7.32E-04	0.003	0.828
AU	SM	6.682	0.463	5.77	0.449	5.58E-04	5.85E-04	0.34
CAD	ADHD	53.422	0.024 *	47.939	0.057	0.012	0.006	0.049 *
CAD	AU	60.739	0.004 *	60.042	0.004 *	3.05E-04	4.85E-04	0.53
CAD	EA	60.738	0.004 *	59.981	0.004 *	0.003	0.005	0.512
CAD	COPD	67.438	8.04E-04 *	67.261	5.80E-04 *	6.37E-04	0.002	0.764
CAD	SM	68.649	5.81E-04 *	60.943	0.003 *	9.64E-04	4.65E-04	0.038 *
EA	ADHD	34.542	0.538	34.235	0.505	-0.003	0.006	0.579
EA	AU	108.685	3.19E-09 *	107.673	2.53E-09 *	-4.42E-04	7.70E-04	0.566
EA	CAD	38.804	0.344	37.161	0.37	2.50E-04	2.01E-04	0.214
EA	COPD	105.393	9.94E-09 *	102.569	1.49E-08 *	-0.003	0.003	0.326
EA	SM	93.471	5.31E-07 *	92.964	3.75E-07 *	2.97E-04	6.81E-04	0.662
COPD	ADHD	11.08	0.197	4.415	0.731	0.031	0.012	0.010 *
COPD	AU	7.242	0.511	5.696	0.576	0.001	8.91E-04	0.214
COPD	CAD	28.47	3.93E-04 *	28.319	1.92E-04 *	-1.54E-04	7.96E-04	0.847
COPD	EA	9.185	0.327	8.252	0.311	0.009	0.01	0.374
COPD	SM	12.911	0.115	11.04	0.137	-0.001	0.001	0.276
SM	ADHD	26.406	7.44E-05 *	9.358	0.053	0.063	0.023	0.007 *
SM	AU	13.45	0.020 *	10.805	0.029 *	-0.002	0.002	0.322
SM	CAD	9.85	0.08	3.502	0.478	0.001	5.06E-04	0.012 *
SM	EA	99.998	5.29E-20 *	84.536	1.90E-17 *	-0.048	0.056	0.392
SM	COPD	11.388	0.044 *	7.531	0.11	-0.009	0.006	0.152

Supplementary Table 6: Robustness measures for multivariable Mendelian randomization. ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, S.E. = standard error, SM = Smoking. x = this value could not be computed due to the low number of single nucleotide polymorphisms (SNPs) in the genetic instrumental variable (IV).

Exposure	Outcome	Conditional F Statistic					Cochran's Q_A		MV MR-Egger Intercept		
		Exposure 1	Exposure 2	Exposure 3	Exposure 4	Exposure 5	Cochran's Q_A	P	Intercept estimate	S.E.	P
ADHD	AU	0.675	1.253	1.023	0.559	0.812	6.003	0.306	-9.17E-04	0.005	0.846
ADHD	CAD	0.679	1.732	0.948	0.544	1.002	4.826	0.437	-8.76E-04	0.002	0.567
ADHD	EA	0.859	1.386	0.903	0.951	0.706	6.63	0.25	1.42E-04	0.085	0.999
ADHD	COPD	0.855	1.231	0.826	3.046	0.696	3.803	0.578	0.012	0.008	0.121
ADHD	SM	0.676	0.961	0.846	0.691	0.433	6.925	0.226	-0.004	0.002	0.025 *
AU	ADHD	0.237	1.179	0.423	0.226	0.186	2.701	0.259	0.011	0.014	0.437
AU	CAD	0.241	0.952	0.471	0.242	0.201	3.745	0.154	-4.25E-04	9.00E-04	0.636
AU	EA	0.261	0.427	0.549	0.196	0.245	5.138	0.077	0.007	0.043	0.869
AU	COPD	1.283	0.486	0.685	1.143	0.828	1.308	0.52	0.002	0.004	0.702
AU	SM	0.657	0.453	0.622	0.934	0.732	1.319	0.517	6.33E-04	7.90E-04	0.423
CAD	ADHD	4.247	1.248	1.359	1.496	1.648	51.993	0.008 *	0.014	0.007	0.034 *
CAD	AU	4.514	1.441	1.396	1.733	1.712	40.93	0.088	4.29E-05	5.27E-04	0.935
CAD	EA	4.862	1.478	1.321	1.625	1.754	47.18	0.024 *	0.002	0.006	0.727
CAD	COPD	5.397	1.449	1.478	1.458	1.648	52.232	0.007 *	9.91E-04	0.002	0.683
CAD	SM	9.36	1.454	1.34	1.443	1.501	54.864	0.004 *	9.97E-04	4.97E-04	0.045 *
EA	ADHD	4.786	2.522	0.992	2.586	2.404	30.21	0.506	-0.006	0.006	0.333
EA	AU	1.458	0.837	1.029	2.552	2.039	82.813	1.32E-06 *	-9.13E-04	8.07E-04	0.258
EA	CAD	1.512	0.833	2.668	2.518	2.033	34.625	0.299	2.97E-04	2.11E-04	0.159
EA	COPD	1.385	0.862	2.623	1.006	2.082	89.527	1.37E-07 *	-0.004	0.003	0.269
EA	SM	1.469	0.913	2.542	1	2.556	68.432	1.22E-04 *	6.72E-04	7.20E-04	0.351
COPD	ADHD	2.675	0.694	2.07	0.704	0.824	2.337	0.505	0.02	0.015	0.194
COPD	AU	2.498	0.877	1.784	0.7	0.884	2.253	0.522	-3.57E-04	0.001	0.812
COPD	CAD	1.031	0.377	0.312	0.665	0.897	7.102	0.069	1.38E-04	0.002	0.936
COPD	EA	1.316	0.383	0.317	2.097	0.986	4.245	0.236	0.004	0.022	0.875
COPD	SM	1.61	0.289	0.261	1.05	0.497	8.599	0.035 *	-5.71E-04	0.003	0.819
SM	ADHD	0.211	0.328	0.292	1.202	0.197	2.882	0 *	0.117	x	x
SM	AU	0.431	0.725	0.394	0.95	0.402	1.394	0 *	-0.016	x	x
SM	CAD	0.494	1.254	0.599	1.406	0.514	1.159	0 *	0.002	x	x
SM	EA	0.172	0.485	0.251	0.214	0.168	7.409	0 *	0.616	x	x
SM	COPD	1.251	0.806	1.4	0.916	1.219	1.001	0 *	0.059	x	x

Supplementary Table 7: The number and percentage of genetic instrumental variable (IV) single nucleotide polymorphisms (SNPs) which were significantly associated with their respective trait in the replication genome-wide association studies (GWASs) for different significance thresholds. ADHD = Attention-Deficit/Hyperactivity Disorder, AU = Alcohol Use, CAD = Coronary Artery Disease, COPD = Chronic Obstructive Pulmonary Disease, EA = Educational Attainment, SM = Smoking.

<i>P</i> value	ADHD		AU		CAD		EA		COPD		SM	
	#	%	#	%	#	%	#	%	#	%	#	%
No threshold	11	100	12	100	41	100	39	100	13	100	10	100
0.05	11	100	8	67	36	88	37	95	9	69	5	50
0.01	11	100	7	58	32	78	36	92	5	38	3	30
5 x 10⁻⁴	11	100	6	50	28	68	31	79	5	38	2	20
5 x 10⁻⁸	10	91	4	33	17	41	23	59	4	31	1	10