

Evaluating the effects of alcohol and tobacco use on cardiovascular disease using multivariable Mendelian randomization

Daniel B. Rosoff¹, George Davey Smith², Nehal Mehta³, Toni-Kim Clarke⁴, Falk W. Lohoff^{1#}

¹Section on Clinical Genomics and Experimental Therapeutics, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD, USA

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

³Section of Inflammation and Cardiometabolic Diseases, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA

⁴Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK

#Corresponding Author:

Falk W. Lohoff, M.D.

Chief, Section on Clinical Genomics and Experimental Therapeutics (CGET)

Lasker Clinical Research Scholar

National Institute on Alcohol Abuse and Alcoholism (NIAAA)

National Institutes of Health

10 Center Drive (10CRC/2-2352)

Bethesda, MD 20892-1540

Office: 301-827-1542

falk.lohoff@nih.gov

ABSTRACT

Alcohol and tobacco use, two major modifiable risk factors for cardiovascular disease (CVD), are often consumed together. Using large publicly available genome-wide association studies (results from > 940,000 participants), we conducted two-sample multivariable Mendelian randomization (MR) to simultaneously assess the independent effects of alcohol and tobacco use on CVD risk factors and events. We found genetic instruments associated with increased alcohol use, controlling for tobacco use, associated with increased high-density-lipoprotein-cholesterol (HDL-C), decreased triglycerides, but not with coronary heart disease (CHD), myocardial infarction (MI), nor stroke; and instruments for increased tobacco use, controlling for alcohol use, associated with decreased HDL-C, increased triglycerides, and increased risk of CHD and MI. Exploratory analysis found associations with HDL-C, LDL-C, and intermediate-density-lipoprotein metabolites. Consistency of results across complementary methods accommodating different MR assumptions strengthened causal inference, providing strong genetic evidence for the causal effects of modifiable lifestyle risk factors on CVD risk.

INTRODUCTION

Cardiovascular disease (CVD) is a leading cause of mortality and morbidity with an estimated annual 17.9 million deaths worldwide¹. Evidence shows that addressing modifiable lifestyle factors can prevent most CVDs². Two highly prevalent and frequently co-occurring risk factors, alcohol consumption and smoking (affecting 2 billion and 1.1 billion people worldwide, respectively^{2,3}) are associated with a considerable proportion of CVD mortality. For example, one in ten CVD deaths are attributable to smoking⁴, and observational studies strongly suggest that smoking increases the incidence of myocardial infarction (MI) and coronary heart disease (CHD)⁵.

Using novel epidemiological methods, the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) recently reported that any level of alcohol monotonically increases risk for many health-related harms, including cancer (GBD)⁶; however, considerable debate remains regarding the association between alcohol consumption and CVD risk^{7,8}.

Observational studies show a complex relationship between alcohol consumption and CVDs with some studies reporting light-to-moderate alcohol use associated with moderately reduced risk of MI⁹ and coronary heart disease^{10,11}, while heavier alcohol consumption and binge drinking are associated with increased stroke, MI, and CHD^{11,12}. Meta-analyses and short-term trials suggest alcohol consumption is associated with CVD risk factors including increased high-density lipoprotein-cholesterol (HDL-C)¹³; however, the association with low-density lipoprotein cholesterol (LDL-C) and triglycerides (TRG) are less clear¹⁴. Variation in size, density, concentrations and

composition of circulating lipids, which traditional measures are unable to distinguish, are thought to have contrasting effects to CVD risk¹⁵. Metabolomics provides a detailed view of systemic metabolism, and the physiological effect of alcohol consumption extends across multiple metabolic pathways⁷ and has been associated with molecular dysregulation conferring both reduced and increased CVD risk⁷.

Despite its public health importance¹³, nearly all human data evaluating the alcohol-CVD relationship is from conventional epidemiological studies¹⁴, and thus, the causal effect of alcohol consumption patterns on CVD risk factors and outcomes has not been fully elucidated. Triangulating evidence from different study designs is needed to improve our understanding of the alcohol-CVD relationship¹⁶. However, observational studies are subject to potential confounding and reverse causation, which makes causal inferences difficult¹⁷ and many of the interventional studies or clinical trials examining the association between alcohol and CVD risk factors were small, failed to randomize different amounts of alcohol, and only observed short-term effects^{13,18}. Further, the recent Moderate Alcohol and Cardiovascular Health (MACH) study, which was a large multicenter, worldwide, randomized clinical trial of ~15 gm of alcohol daily versus abstinence, has been stopped, due to concerns about study design and future credibility¹⁹.

One alternative strategy to investigate potential causal inference between an exposure and an outcome, in particular when randomized control trials (RCTs) are not feasible, practical, or ethical, would be Mendelian randomization analysis (MR)²⁰. MR analyses can be conceptualized as natural RCTs²¹. MR utilizes randomly inherited genetic

variants, established before outcome onset and therefore relatively independent of confounders, as instruments for an exposure to assess the causal effect of the risk factor exposure on a health outcome of interest²². Where it may be difficult to find a genetic variant uniquely associated with a risk factor, MR has been extended to multivariable MR (MVMR) using genetic variants associated with multiple risk factors—alcohol and tobacco use in this study—to simultaneously and consistently estimate the effect of each of the risk factors on the outcome(s) of interest^{23,24}. Assumptions underlying single and multivariable MR analysis require sufficient statistical power to perform the required sensitivity analyses²². The two-sample single- and multivariable MR analysis used in this study—using summary-level data from separate genome-wide association studies (GWASs) for exposures and outcomes—allows for the use of complementary MR methods capable of investigating these assumptions²⁵.

Previous one-sample MR studies using genetic variation in either the acetaldehyde dehydrogenase (*ALDH*) or alcohol dehydrogenase (*ADH*) genes found alcohol associated with increased HDL-C levels^{8,18,26}. MR studies have shown that alcohol consumption reduces triglycerides and CHD risk⁸. A recent MR study using 512,715 Chinese adults from the Kadoori Biobank found alcohol increased stroke risk²⁷, but had no effect on MI, possibly due to low MI prevalence in the sample²⁷. Given the complexity of the relationship between alcohol and CVDs, average alcohol consumption may not be sufficient to evaluate the risk conferred for CVDs²⁸. There is observational evidence for a link with approximately 85% of smokers drinking alcohol²⁹, and drinkers are 75% more likely than abstainers to smoke³⁰, as well as a strong genetic correlation^{31,32}. Thus, failure

to account for smoking in previous single variable MR analyses (SVMR) of alcohol and CVDs may have produced biased and inconsistent estimates.

Using a two-sample MR design (Fig. 1), we created instruments for alcohol and tobacco use using summary statistics from publicly available, large-scale genome-wide association studies (GWAS)³² (Table 1 and Supplementary Table 1) to evaluate the potential causal effects thereof on cardiovascular disease risk factors and disease events. Given the strong association between drinking and smoking³³, we conducted multivariable MR analyses to estimate the direct effect of each of exposures—alcohol and tobacco use—controlled for the other—on our cardiovascular outcomes of interest. We compared these direct independent effects to the total effects estimated by single variable MR analysis. We also conducted exploratory multivariable MR analyses examining how alcohol and tobacco use effected changes in the metabolomic profile: assessing the detailed metabolomic effects of alcohol and tobacco use could inform our understanding of association thereof with lipid CVD biomarkers. Complementary MR techniques—inverse-variance weighted MR and MR-Egger—were used as sensitivity analyses to test the robustness of all results.

RESULTS

Effects of alcohol and tobacco use on cardiovascular disease lipid risk factors and events. In SVMR, we found genetic variants associated with increased alcohol use, not controlling for tobacco use, associated overall with increased levels of HDL-C (β , 0.213, 95% CI, 0.101- 0.325, $P= 2.03 \times 10^{-4}$), decreased levels of TRG (β , -0.465, 95% CI, -

0.616- -0.314, $P= 1.33 \times 10^{-9}$), but not with levels of LDL-C nor TC. These results are shown in Fig. 2 and Table 2 in the Supplement. In MVMR, we found these variants, controlling for effects of tobacco use, associated with increased levels of HDL-C and decreased levels of TRG, but to a lesser extent (DPW: β , 0.195, 95% CI, 0.053- 0.091, $P= 0.000$; TRG: β , -0.205, 95% CI, - 0.338, -0.072, $P= 0.002$). These results are shown in Fig. 2 and Table 3 in the Supplement. However, we did not find these variants associated, in either SVMR or MVMR, with risk of coronary heart disease (CHD), myocardial infarction (MI), nor any stroke, including ischemic, cardioembolic, large artery, and small vessel stroke.

In SVMR, we found genetic variants associated with increased risk of tobacco use (ever smoker status) associated, not controlling for alcohol use, with decreased levels of HDL-C (β , -0.145, 95% CI, -0.218- -0.072, $P= 1.15 \times 10^{-4}$), increased levels of TRG (β , 0.137, 95% CI, 0.057 -0.217, $P= 8.31 \times 10^{-4}$), but not with levels of LDL-C nor TC. We also found these variants associated with increased risk of CHD (OR, 1.309, 95% CI, 1.143- 1.498, $P= 1.05 \times 10^{-5}$) and MI (OR, 1.48, 95% CI, 1.25-1.752, $P= 4.88 \times 10^{-6}$). We did not find these variants associated with any stroke, nor with ischemic, cardioembolic, and small vessel stroke. These results are shown in Fig. 2 and Table 4 in the Supplement. In MVMR, we found these variants, controlling for alcohol use, associated with decreased levels of HDL-C (β , -0.159, 95% CI, -0.083- -0.235, $P= 1.15 \times 10^{-4}$), increased levels of TRG (β , 0.124, 95% CI, 0.034 -0.214, $P= 8.31 \times 10^{-4}$), although again to a lesser extent than in SVMR. We found these variants, controlling for alcohol use, associated with increased risk of CHD (OR, 1.368, 95% CI, 1.185-1.578, $P= 0.0000$) and MI (OR, 1.48,

95% CI, 1.193-1.685, $P= 1.05 \times 10^{-4}$). These results are shown in Fig. 2 and Table 3 in the Supplement. We did not find these variants associated with any stroke, nor with ischemic, cardioembolic, and small vessel stroke, under either SVMR or MVMR.

MR-Egger can detect, under certain assumptions, directional pleiotropy; MR-Egger estimates are robust to such pleiotropy. While the magnitude and direction of the MR IVW effect estimates were not substantially different under SVMR and MVMR, the MR-Egger estimates were more precise under MVMR, simultaneously estimating the effects of alcohol and tobacco use. Further, the MR IVW effect estimates were more consistent (similar in magnitude and direction) with these MR-Egger estimates under MVMR analysis, strengthening our hand inferring causality from the MVMR results (Fig. 2 and Tables 2-4 in the Supplement).

Effects of alcohol and tobacco use on lipoprotein and lipid metabolites. In

exploratory MVMR analyses (reported here given our main interest in the direct versus the total (direct plus indirect) effects of alcohol and tobacco use), we found genetic variants associated with increased alcohol use, controlling for tobacco use, associated with increased concentrations of IDL particles (β , 0.253, 95% CI, 0.065-0.441, $P= 0.009$), large LDL particles (β , 0.280, 95% CI, 0.088-0.472, $P= 0.004$), medium LDL particles (β , 0.253, 95% CI, 0.061-0.445, $P= 0.012$), small LDL particles (β , 0.267, 95% CI, 0.069-0.4465, $P= 0.008$); very large HDL particles (β , 0.285, 95% CI, 0.087-0.483, $P= 0.002$), large HDL particles (β , 0.396, 95% CI, 0.218-0.574, $P= 0.000$), and medium HDL particles (β , 0.248, 95% CI, 0.072-0.424, $P= .011$), but not with small HDL

particles. These results are shown in Fig. 3, column 1, and Table 5 in the Supplement.

These alcohol use variants were further associated with increased levels of total cholesterol in IDL (β , 0.309, 95% CI, 0.123-0.495, $P= 0.001$), small LDL (β , 0.280, 95% CI, 0.086-0.474, $P= 0.005$), medium LDL (β , 0.242, 95% CI, 0.052-0.432, $P= 0.013$), LDL (β , 0.275, 95% CI, 0.087-0.463, $P= 0.004$) and large LDL (β , 0.256, 95% CI, 0.070-0.442, $P= 0.007$); also HDL (β , 0.310, 95% CI, 0.145-0.475, $P= 2.43 \times 10^{-4}$), large HDL (β , 0.333, 95% CI, 0.171-0.495, $P= 5.81 \times 10^{-5}$), and very large HDL (β , 0.229, 95% CI, 0.053-0.405, $P= 0.011$); not with levels of free cholesterol, but with free cholesterol to esterified cholesterol ratio (β , 0.371, 95% CI, 0.126-0.616, $P= 0.003$). We did not find these variants associated with any triglyceride metabolites, except for triglycerides in very large HDL (β , 0.230, 95% CI, 0.046-0.414, $P= 0.014$). These results are shown in Fig. 3, column 1, and Table 5 in the Supplement.

Further still, these alcohol use variants were associated with increased levels of phospholipids in IDL (β , 0.277, 95% CI, 0.101-0.473, $P= 0.003$); medium LDL (β , 0.233, 95% CI, 0.047-0.419, $P= 0.014$); large LDL (β , 0.248, 95% CI, 0.062-0.434, $P= 0.009$); and medium HDL (β , 0.242, 95% CI, 0.060-0.424, $P= 0.009$), large HDL (β , 0.405, 95% CI, 0.229-0.581, $P= 4.90 \times 10^{-6}$), and very large HDL (β , 0.338, 95% CI, 0.162-0.514, $P= 4.90 \times 10^{-6}$); as well as increased levels of apolipoprotein B (β , 0.242, 95% CI, 0.048-0.436, $P= 0.14$) and A-1 (β , 0.300, 95% CI, 0.116-0.484, $P= 0.001$). These results are shown in Fig. 3, column 1, and Table 5 in the Supplement.

Genetic variants associated with risk of tobacco use, controlling for alcohol use, were associated with fewer metabolites than, and in the direction opposite the alcohol use variants: We found tobacco use variants associated with decreased concentrations of large LDL particles (β , -0.167, 95% CI, -0.320--0.014, $P= 0.032$), medium LDL particles (β , -0.177, 95% CI, -0.334--0.020, $P= 0.026$), and small LDL particles (β , -0.173, 95% CI, -0.330--0.016, $P= 0.031$); cholesterol in IDL (β , -0.172, 95% CI, -0.319--0.025, $P= 0.022$), small LDL (β , -0.163, 95% CI, -0.314--0.012, $P= 0.040$); medium LDL (β , -0.145, 95% CI, -0.294--0.004, $P= 0.056$), LDL (β , -0.154, 95% CI, -0.301--0.007, $P= 0.026$), and large LDL (β , -0.168, 95% CI, -0.313--0.023, $P= 0.023$), and also free cholesterol to esterified cholesterol ratio (β , -0.260, 95% CI, -0.452--0.068, $P= 0.008$). We also found tobacco use variants associated with decreased levels of phospholipids, but only in medium LDL (β , -0.153, 95% CI, -0.298--0.008, $P= 0.040$) and large LDL (β , -0.152, 95% CI, -0.297--0.007, $P= 0.041$), but not in any triglycerides or apolipoproteins. These results are shown in Fig. 3, column 2, and Table 5 in the Supplement.

DISCUSSION

Our two-sample MVMR design allowed us to estimate the direct effects of alcohol and tobacco use on cardiovascular disease risk factors and certain risk events, as opposed to and distinguished from the total effects estimated by single variable design. Specifically, total alcohol use was associated with increased levels of HDL-C, but decreased levels of triglycerides. Extending traditional lipid measures, our study also provides a comprehensive examination of the associations of alcohol and tobacco use on lipid particles and lipid constituents. In general, our findings support previous observational

research showing a relationship between alcohol, smoking, and CVDs^{5,13}, and the consistent estimates across complementary MR methods and after controlling alcohol consumption for smoking (and smoking for alcohol consumption) strengthen causal inference. Finding robust increases in HDL-C with no corresponding reduction in CVD risk calls into question the mechanism through which observational literature attributed any cardioprotective effect of light-to-moderate alcohol consumption¹⁴. Similarly, our results, when combined with traditional RCTs and MR studies that failed to find CVD risk reduction of increased HDL-C³⁴⁻³⁶ suggests that there is no cardioprotective effect of alcohol consumption.

Importantly, in our main analysis we failed to find a cardioprotective effect of alcohol use on CVD outcomes, suggesting that previously reported reductions in CVD risk due to other lifestyle differences that are associated with more light-to-moderate drinking patterns, such as healthier lifestyles³⁷, are contributing to these effects. Our total weekly consumption instrument may have failed to differentiate between light-to-moderate drinking and heavier consumption, which are reported to have inverse associations with CVD disease risk in observational literature¹⁴. More likely, however, previous research may not have accounted for smoking, which we found to be significantly associated with increased CVD risk. For example, we failed to replicate an early MR study that found carriers of the *ADH1B* rs1229984 A-allele (associated with reduced alcohol consumption) were also associated with lower CHD and stroke risk⁸. However, they reported that rs1229984 was also associated with smoking: omitting smoking from the alcohol consumption equation would have influenced their results. They also found rs1229984

was associated with increased educational attainment, which more recent MR studies have shown to be causally linked with CVD risk³⁸. Expanding upon their single stroke variable, which may mask differential associations between alcohol and stroke subtypes¹¹ further strengthens our conclusions. Failing to replicate the instrumental variable estimates Millwood et al. 2019 found among participants in the China Kadoorie Biobank sample supports the hypothesized differential effect of alcohol consumption on CVD risk across ethnic backgrounds^{14,39}. East Asia has the highest prevalence of stroke worldwide⁴⁰, and generally lower alcohol consumption behaviors than European countries³. In addition, survey data show that men consume 95% of all alcohol in China⁴¹. Therefore, caution is needed before extrapolating these findings to other populations and our study highlights the need for population-specific MR analysis in elucidating the underlying alcohol-CVD relationship.

Consistent with results derived from observational literature, we found strong genetic evidence of a causal effect of tobacco use, independent of alcohol use, on increased MI and CHD risk⁵. Our estimates showing 48% and 30% increased risk for MI and CHD are similar to estimates from observational and smoking cessation studies⁴², and we confirm findings from MR study that found a smoking instrument constructed from multiple smoking behaviors increased CHD and MI risk³⁸ by replicating their results using a separate smoking instrument accounting for alcohol consumption. Mechanisms underlying the association with MI may include nicotine-induced sympathetic nervous system activation and vascular endothelial function impairment, which together may induce cardiovascular stress through increased blood pressure, heart rate, and cardiac

output⁴³, while long-term effects on atherosclerotic plaques, thrombosis, and inflammatory processes may mediate CHD risk⁵. While an increased awareness and understanding of the adverse health effects associated with smoking has decreased the prevalence of current smoking, smoking continues to be a major cardiovascular hazard⁴³. For example, in the United States, while the percentage of adults reporting having ever smoked has decreased from 43% in 1965 to 19% in 2010⁴³, 44 million adults still smoke. We were unable to assess the causal effect of smoking on CVD mortality; however, a meta-analysis found smoking cessation associated with a 36% reduced mortality risk, which is similar in magnitude to other prevention strategies, including statins⁴⁴, and β -blockers⁴⁴. If the association is causal, our results may be useful to further inform continued clinical trial and smoking prevention program development.

We failed to replicate the increase in stroke risk associated with smoking Carter et al. 2019 found using a constructed smoking instrument among UK Biobank and MEGASTROKE participants³⁸. Our smoking instrument was based upon the “ever smoker status”, selected to pick up generalized lifetime smoking behavior for our primary alcohol use analysis; to the extent the “smoker” status included current “non-smokers”, we would expect some impact on the effect of smoking on risk of stroke risk. Given the strong evidence of a substantial adverse effect of even one cigarette per day on risk of CVD and stroke⁴⁵, caution is needed when interpreting our null finding.

Our findings also support the causal role of alcohol consumption in increasing HDL-C as reported in previous MR studies^{18,26} and provides evidence for the finding after controlling for smoking behavior. While we were unable to evaluate the effects of heavy

alcohol consumption on HDL-C levels, we recently reported a cross-sectional study findings that very heavy alcohol consumption (> 10 standard drinks for men and > 8 standard drinks for women) to be associated with increased HDL-C⁴⁶, which suggests that alcohol may impact HDL-C levels across a range of alcohol consumption levels. Future MR studies are needed to investigate this hypothesis. Our comprehensive examination of the associations of alcohol with lipid particles and lipid constituents replicated a recent metabolomic study⁷, and revealed that alcohol may increase large HDL-C particle size, total cholesterol concentration and apolipoprotein A1, which each are associated with reduced MI risk¹⁵; however, when considering coincident increases in MI increasing constituents such as small and medium LDL-C and large LDL-C particle size¹⁵, and, as mentioned above, the RCTs and MR studies failing to find increased HDL-C and reduced CVD risk^{8,34-36}, further suggests any hypothetical cardioprotective mechanism derived from HDL-C metabolism changes may be further reduced when compared to assessments failing to consider these heterogenous associations. It is possible that alcohol consumption may reduce CHD risk through its modulation of HDL-C phospholipid content, which impacts reverse cholesterol transport⁴⁷ – a pathways inversely associated with prevalent CHD, even after controlling for HDL-C levels⁴⁸ – and has been shown to be reduced among CHD patients⁴⁷.

Our finding that alcohol use reduces TRG confirms results from MR studies using European^{8,18} and European American²⁶ cohorts; however, our 58 SNP instrument for alcohol use may account, in part, for the reported direction. The association between alcohol and triglycerides remains equivocal¹⁴ with a meta-analysis of intervention studies reporting no association¹³, and observational studies have reported alcohol associated

with either reduced, or harmful TRG levels, depending on the cohort and the amount consumed⁴⁹. However, the intervention studies investigated primarily short-term effects while genetic variants associated with alcohol consumption are constant throughout the life course¹⁸. Since our finding was also in a European American sample, and alcohol may affect TRG levels differently depending on race or ethnicity⁵⁰, future trans-ancestry replication studies are needed.

Supporting the observational literature⁵, we found genetic evidence for a causal effect of smoking on decreased levels of HDL-C and increased levels of TRG. Conversely, our findings do not support Åsvold et al. 2014's MR analysis of the Norwegian HUNT Study cohort showing higher HDL-C and no effect on triglycerides using a single SNP instrument (rs1051730 located in *CHRNA3*)⁵¹. However, associations between rs1051730 and HDL-C were not significantly different by smoking status in the HUNT cohort, and even the authors cautioned interpretation of their HDL-C finding⁵¹. Smoking has been shown to alter lipid metabolism enzymes, which would alter HDL-C levels⁵². Further, HDL-C particles can be rendered nonfunctional due to oxidative modifications from smoking, which suggests smoking may impact HDL-C quantity and function⁵².

Our study has several strengths. We use a multivariable two-sample MR design, the most appropriate design given the strong correlation between alcohol and tobacco use, yielding unbiased and consistent estimates of the direct effects of our multiple exposures on the CVD risk factors and disease events²⁴. We selected the SNPs included in the genetic instruments through a stringent process. Another strength is the use of large publicly available summary-level genetic study datasets (from more than 940,000 study

participants overall) that are necessary for two-sample MR⁵³ and the broadly consistent estimates across multiple MR methods, each relying on orthogonal assumptions, provides confidence in result robustness and strengthens causal inference²⁵. We also found no evidence of horizontal pleiotropy, suggesting minimal potential bias²². Finally, while previous MR studies only used a combined stroke variable, we used multiple stroke subtypes, given some observational studies reporting a differential association of alcohol by stroke subtype¹¹, which would strengthen our conclusion regarding a null effect of alcohol use on stroke risk. Similarly, our metabolomic analysis highlighting the complex impact of alcohol use on circulating metabolites may be an important addition to the field's growing body of MR literature.

Several limitations should be noted. As with previous alcohol literature, alcohol consumption patterns are variable, and the exposure variables may be either under- or over-reported²⁷. Carter et al. 2019 found that smoking mediates the relationship between education and CVD risk³⁸, and we recently reported that a two-sample MR finding educational attainment (EA) affects how alcohol is consumed – with increased EA increasing whether alcohol is consumed with meals, drinking more frequently, but less likely to binge drink (≥ 6 drinks per occasion)⁵⁴. Considering previous work has shown that how alcohol is consumed is associated with different dietary patterns⁵⁵, and differences in CVD risk²⁸, it is possible dietary patterns may moderate the association between alcohol and CVD risk. Therefore, future studies using cohorts with heavier reported consumption or incorporating dietary patterns associated with different drinking typologies would further elucidate the alcohol-CVD relationship.

While genetic instruments are considered to represent lifelong exposure²², currently no GWAS available examines the genetic architecture of alcohol consumption throughout the entire life course. In addition, while we show an effect of alcohol consumption and smoking on certain CVD risk factors, alcohol has been associated with other biological mechanisms, such as, such as cellular adhesion molecules and adipocyte hormones, that may contribute to CVD risk^{13,14}, which we did not assess. We examined total alcohol consumption among participants in the UK Biobank, who have been reported to be more educated, have healthier lifestyles, and fewer health problems than the UK population⁵⁶, and other hazardous consumption patterns associated with increased CVD risk⁵⁷, were not assessed. .

Further, the included GWAS are from survey participants of European ancestry, and given observational evidence suggesting different associations between alcohol and CVDs exist across racial and ethnic groups¹⁴, caution may be advised in extrapolating these findings to other populations. Genome-wide significant (GWS) SNPs for each exposure only explain a small fraction of total variance in complex traits⁵⁸. Relatedly, some GWS exposure SNPs, or their proxies, were not available in every outcome GWAS used in this study, suggesting reduced variance captured with the remaining SNPs²⁵. Also, given the absence of available outcome GWASs, we were unable to replicate the findings in independent datasets. While we found no evidence of pleiotropy, it is possible pleiotropy still exists from unmeasured confounders. Relatedly, while unlikely, we cannot entirely rule out the possibility of sample overlap between exposure and

outcome datasets, which would bias estimates⁵⁹; however, any bias would likely be minimal⁵⁹.

Using two-sample multivariable MR, we show a likely direct causal effect of alcohol consumption on increased HDL-C, and we failed to find a cardioprotective effect of alcohol. However, smoking significantly increased the risk for MI, CHD, and stroke. Our findings suggest previous observational studies reporting an association between alcohol consumption and CVDs may have resulted from confounding, and differences between our findings and recent MR analyses using data from populations with different ethnic background suggest that future MR studies in additional populations may help elucidate this complex relationship. Our exploratory metabolomic analyses highlight the complex impact of alcohol on both cardioprotective biomarkers and biomarkers associated with increased CVD risk. Our findings may help inform the public health debate regarding the alcohol-CVD relationship and suggest that future work examining this relationship must account for smoking behaviors. In conjunction with other study designs, this study provides additional evidence that programs targeted at reducing alcohol consumption and smoking may have beneficial cardiovascular effects.

METHODS

Data sources. We conducted single-variable (SV) as well as multivariable (MV) two-sample MR analyses of alcohol and tobacco use on cardiovascular disease and disease risk factors using summary-level data from publicly available GWASs from similar populations of predominantly European ancestry of each of alcohol and tobacco use and

cardiovascular disease and disease risk factors (Table 1 and Table 1 in the Supplement).

All studies have existing ethical permissions from their respective institutional review boards and include participant informed consent and included rigorous quality control (Methods 1 in the Supplement). Figure 1 summarizes the study design.

Instruments. SNPs for alcohol use were extracted from a large GWAS meta-analysis of 414,343 individuals of European ancestry from the UK Biobank prospective cohort study collected across the United Kingdom from 2006-2010. We included all SNPs associated at genome-wide significance ($P < 5 \times 10^{-8}$) (we use the standard heuristic significance threshold for these dimensional MR to make the work manageable) with the phenotype “drinks per week” (DPW), constructed, for UK Biobank participants who indicated they drank “at least once or twice per week”, by aggregating the weekly intake of distilled spirits (pub measures), beer and cider (pints), red wine, white wine, and champagne (glasses), and other alcoholic drinks *e.g.* alcopops (DPW: mean = 8.92 drinks, SD = 9.30 drinks)³². (For UK Biobank participants who indicated they drank “one to three times a month”, the phenotype was constructed by aggregating the monthly intake over all drink types and dividing by four.) We pruned the results to exclude all SNPs with a pairwise linkage disequilibrium (LD) $R^2 > 0.001$, giving us 58 independent SNPs associated with alcohol use for the single-variable MR (SVMR) analyses (Table 6 in the Supplement).

For tobacco use, we extracted SNPs from the companion GWAS meta-analysis of 518,633 individuals of predominantly European ancestry (444,598 from the UK Biobank prospective cohort study and 74 035 from the Tobacco, Alcohol and Genetics (TAG)

Consortium)³². We included all SNPs associated at genome-wide significance with the phenotype “ever smoker status” (a binary phenotype coded “1” if the individual reported he or she was a previous or current smoker, “0” if he or she had never smoked or only smoked once or twice). We pruned the results to exclude all SNPs with a pairwise LD $R^2 > .001$, giving us 141 independent SNPs associated with tobacco use for SVMR analyses (Table 7 in the Supplement).

For the multivariable MR (MVMR) analyses, we included all 199 SNPs which were genome-wide significant in either the GWASs of either alcohol or tobacco use, deleting 2 duplicate SNPs (SNPs associated at genome-wide significance with both alcohol and tobacco use). (We note as regards number of unique instruments for each exposure, that 14 of the 58 instruments for alcohol use were associated with tobacco use (at nominal P of 0.05 corrected for 58 comparisons); 16 of the 141 instruments for tobacco use were associated with alcohol use (at nominal P of 0.05 corrected for 141 comparisons)). We pruned the results to exclude any SNPs with a pairwise LD $R^2 > 0.001$, giving us 169 independent SNPs for the MVMR analyses.

Cardiovascular risk factors. We used summary statistics from the GWAS meta-analyses for lipid levels – HDL-C, LDL-C, TC and triglycerides (measured in SD units (mg/dL)) – in 37 cohorts (none UKB nor TAG) comprising 188,577 persons of predominantly European ancestry (excluding persons known to be on lipid lowering medications or women known to be pregnant)⁶⁰. Of the 58 possible genome-wide significant SNPs associated with alcohol use, 28 SNPs were present in the GWASs, 30

SNPs were missing, 14 SNPs were identified in high linkage disequilibrium (LD) as proxies for some of the missing SNPs were present, 5 SNPs were removed during harmonization for being palindromic with intermediate allele frequency, leaving 37 SNPs for SVMR analysis. Of the 141 possible SNPs associated with tobacco use, 98 SNPs were present in the GWASs, 43 were missing, 23 SNPs were identified in high linkage disequilibrium (LD) as proxies for some of the missing SNPs were present, and 11 SNPs were removed during harmonization for being palindromic with intermediate allele frequency, leaving 110 SNPs for SVMR analysis. Of the 169 possible SNPs of the combined instrument set, 110 SNPs were present in the GWASs, 59 were missing, 27 SNPs were identified in high linkage disequilibrium (LD) as proxies for some of the missing SNPs were present, and 14 SNPs were removed during harmonization for being palindromic with intermediate allele frequency, leaving 123 SNPs for MVMR analysis.

Cardiovascular disease events. We used summary statistics from the GWAS meta-analyses for coronary artery disease (CAD: 60,801 cases; 123,504 controls) and myocardial infarction (CAD sub-group: 43 676 cases; controls 128,199 controls) in 48 cohorts (none UKB nor TAG) in up to 184 305 persons of mixed ancestry (approximately 15% non-European ancestry, including Chinese, Indian Asian, South Asian, Lebanese, African-American and Hispanic American ancestry)⁶¹. Of the 58 possible SNPs associated with alcohol use, all SNPs were present in the GWASs, but 1 was removed during harmonization as palindromic, leaving 57 for SVMR analysis. Of the 141 possible SNPs associated with tobacco use, 140 SNPs were present in the GWASs, and 6 SNPs were removed during harmonization as palindromic with intermediate allele

frequency, leaving 134 SNPs for SVMR analysis. Of the 169 possible SNPs of the combined instrument set, 167 SNPs were present in the GWASs, 2 were missing, 1 SNP was identified in high linkage disequilibrium (LD) as proxies for some of the missing SNPs were present, and 6 SNPs were removed during harmonization for being palindromic with intermediate allele frequency, leaving 162 SNPs for MVMR analysis.

For ischemic stroke and ischemic stroke sub-types, we used summary statistics from the GWAS meta-analyses in 12 cohorts in up to 29,633 persons of European ancestry^{62,63}. Of the 58 possible SNPs associated with alcohol use, 28 SNPs were present in the GWASs, 14 SNPs were identified in high linkage disequilibrium (LD) as proxies for some of the missing SNPs, and 1 SNP was removed during harmonization for being palindromic with intermediate allele frequency, leaving 41 SNPs for SVMR analysis. Of the 141 possible SNPs associated with tobacco use (ever smoker status), 100 SNPs were present in the GWASs, 20 SNPs were identified in high linkage disequilibrium (LD) as proxies for some of the missing SNPs, and 5 SNPs were removed during harmonization as palindromic with intermediate allele frequency, leaving a total of 116 SNPs for analysis. Of the 169 possible SNPs of the combined instrument set, 112 SNPs were present in the GWASs, 57 were missing, 26 SNPs were identified in high linkage disequilibrium (LD) as proxies for some of the missing SNPs were present, and 5 SNPs were removed during harmonization for being palindromic with intermediate allele frequency, leaving 133 SNPs for MVMR analysis.

Circulating metabolites. We used summary statistics from the GWAS meta-analyses of

14 cohorts (none UKB nor TAG) in up to 24,925 persons of European ancestry for 58 phenotypes describing circulating metabolites (under 5 lipoprotein sub-groups headings *i.e.* particle concentrations, triglycerides, cholesterol, phospholipids and apolipoproteins (see Table 1 in the Supplement) ⁶⁴. Of the 169 possible SNPs of the combined instrument set for exploratory MVMR, 162 SNPs were present in the GWASs, 7 were missing, 6 SNPs were identified in high linkage disequilibrium (LD) as proxies for some of the missing SNPs were present, and 6 SNPs were removed during harmonization for being palindromic with intermediate allele frequency, leaving 162 SNPs for MVMR analysis.

Sample overlap. Participant overlap in samples used to estimate genetic associations between exposure and outcome in two sample MR can bias results⁵⁹. We endeavored to use only non-overlapping GWAS summary statistics to reduce this source of weak instrument bias; we avoided overlap between the primary exposure and outcomes cohorts included.

Statistical analysis. For SVMR, we used inverse-variance weighted MR (MR IVW) (single variable weighted linear regression) along with the complementary MR-Egger method to assess the evidence of the causal effects of each of alcohol and tobacco use on cardiovascular disease risk factors and disease outcomes and to detect the sensitivity of the results to different patterns of violations of IV assumptions²²: consistency of results across methods strengthens an inference of causality²². For MVMR, introduced by Burgess et al.²³, we used an extension of the inverse-variance weighted MR method, performing multivariable weighted linear regression (variants uncorrelated, random-

effect model) with the intercept term set to zero⁶⁵. We also used an extension of the MR-Egger method to correct for both measured and unmeasured pleiotropy⁶⁶.

Sensitivity analyses. To evaluate heterogeneity in instrument effects, which may indicate potential violations of the IV assumptions underlying two-sample MR⁶⁷, we used the MR Egger intercept test⁶⁷, and the Cochran heterogeneity test⁶⁸, and multivariable extensions thereof^{65,66}. The MR pleiotropy residual sum and outlier (MR-PRESSO) global test, and multivariable extension thereof⁶⁹, were used to facilitate identification and removal of outlier variants in order to correct potential directional horizontal pleiotropy and resolve detected heterogeneity. Details about MR methods and tests used are included in Methods 2 in the Supplement. For the SVMR, we used the Steiger directionality test to test the causal direction between the hypothesized exposure and outcomes⁷⁰. Given a nominal threshold of .05, with 11 comparisons in the primary analysis, the Bonferroni corrected threshold for the study would be .00454. Analyses were carried out using MendelianRandomization, version 0.4.1, TwoSampleMR, version 4.16²², and MR-PRESSO, version 1.0⁶⁹, in the R environment, version 3.5.1.

Reported results. We generally look for those estimates (1) substantially agreeing in direction and magnitude across MR methods, (2) exceeding nominal significance corrected for ten (10) multiple comparison ($P < 0.005$) in MR IVW, (3) not indicating bias from horizontal pleiotropy (MR-PRESSO global $P > 0.05$, also MR Egger intercept $P > 0.05$), and for SVMR (4) indicating true causal effect directionality (Steiger directionality test $P < 0.05$). Full MR results with test statistics both before and after

outlier correction are presented in Tables 2-5 in the Supplement. MR-PRESSO outlier corrected results are displayed in Figures 2 and 3. Causal effects estimates are presented as the MR IVW (with MR-Egger as the complementary method presented to draw inferences about causality) effect estimate (β) (with 95% confidence interval (CI)), or, for the binary cardiovascular artery disease outcomes, as the odds ratios (OR) (with 95% CI), per unit increase in the exposures, *i.e.* one- unit increase in drinks per week (DPW); or a one-unit increase in the logarithm of the odds ratio ($\log(\text{OR})$) of the risk of ever smoking tobacco (ever smoker status).

ACKNOWLEDGEMENTS

This research was facilitated by the Social Science Genetic Association Consortium (SSGAC), the Substance Use Disorders Working Group of the Psychiatric Genomics Consortium (PGC-SUD) (supported by funds from NIDA and NIMH to MH109532 and, previously, with analyst support from NIAAA to U01AA008401 (COGA)), and the Medical Research Council Integrative Epidemiology Unit (MRC-IEU, University of Bristol, UK), especially the developers of the MRC-IEU UK Biobank GWAS Pipeline. We gratefully acknowledge their contributing studies and the participants in those studies without whom this effort would not be possible. This work was supported by the National Institutes of Health (NIH) intramural funding [ZIA-AA000242 to F.W.L]; Division of Intramural Clinical and Biological Research of the National Institute on Alcohol Abuse and Alcoholism (NIAAA). The authors declare no conflict of interest.

DATA AVAILABILITY

All analyses were conducted using publicly available data. The summary-level data for alcohol and tobacco use are available at <https://www.thessgac.org/data>, and the summary-level data for lipids (GLGC), coronary heart disease (CARDIoGRAMplus4CD), stroke (ISGC), and metabolites are all available through MR Base at <http://www.mrbase.org/>

CODE AVAILABILITY

The analysis code in R is available on request and all data displayed in the figures are available in the Supplementary Tables.

REFERENCES

- 1 Roth, G. A. *et al.* Demographic and Epidemiologic Drivers of Global Cardiovascular Mortality. *New England Journal of Medicine* **372**, 1333-1341, doi:10.1056/NEJMoa1406656 (2015).
- 2 Organization, W. H. (World Health organization Geneva 2018).
- 3 World Health Organization. Management of Substance Abuse Team. *Global status report on alcohol and health*. (World Health Organization, 2011).
- 4 Ezzati, M. *et al.* in *Global Burden of Disease and Risk Factors* (eds A. D. Lopez *et al.*) (2006).
- 5 Ambrose, J. A. & Barua, R. S. The pathophysiology of cigarette smoking and cardiovascular disease: an update. *Journal of the American College of Cardiology* **43**, 1731-1737, doi:10.1016/j.jacc.2003.12.047 (2004).
- 6 Griswold, M. G. *et al.* Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet* **392**, 1015-1035, doi:[https://doi.org/10.1016/S0140-6736\(18\)31310-2](https://doi.org/10.1016/S0140-6736(18)31310-2) (2018).
- 7 Wurtz, P. *et al.* Metabolic profiling of alcohol consumption in 9778 young adults. *Int J Epidemiol* **45**, 1493-1506, doi:10.1093/ije/dyw175 (2016).
- 8 Holmes, M. V. *et al.* Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. *BMJ : British Medical Journal* **349**, g4164, doi:10.1136/bmj.g4164 (2014).
- 9 Wood, A. M., Kaptoge, S. & Butterworth, A. S. Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies (vol 391, pg 1513, 2018). *Lancet* **391**, 2212-2212 (2018).
- 10 Zhao, J., Stockwell, T., Roemer, A., Naimi, T. & Chikritzhs, T. Alcohol Consumption and Mortality From Coronary Heart Disease: An Updated Meta-

- Analysis of Cohort Studies. *J Stud Alcohol Drugs* **78**, 375-386, doi:10.15288/jsad.2017.78.375 (2017).
- 11 Ronksley, P. E., Brien, S. E., Turner, B. J., Mukamal, K. J. & Ghali, W. A. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. *BMJ* **342**, d671, doi:10.1136/bmj.d671 (2011).
- 12 Mukamal, K. J., Maclure, M., Muller, J. E., Sherwood, J. B. & Mittleman, M. A. Prior alcohol consumption and mortality following acute myocardial infarction. *Jama* **285**, 1965-1970 (2001).
- 13 Brien, S. E., Ronksley, P. E., Turner, B. J., Mukamal, K. J. & Ghali, W. A. Effect of alcohol consumption on biological markers associated with risk of coronary heart disease: systematic review and meta-analysis of interventional studies. *BMJ* **342**, d636, doi:10.1136/bmj.d636 (2011).
- 14 Piano, M. R. Alcohol's Effects on the Cardiovascular System. *Alcohol Res* **38**, 219-241 (2017).
- 15 Holmes, M. V. *et al.* Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *Journal of the American College of Cardiology* **71**, 620-632, doi:10.1016/j.jacc.2017.12.006 (2018).
- 16 Lawlor, D. A., Tilling, K. & Davey Smith, G. Triangulation in aetiological epidemiology. *Int J Epidemiol* **45**, 1866-1886, doi:10.1093/ije/dyw314 (2017).
- 17 Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N. & Davey Smith, G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med* **27**, 1133-1163, doi:10.1002/sim.3034 (2008).
- 18 Lawlor, D. A. *et al.* Exploring causal associations between alcohol and coronary heart disease risk factors: findings from a Mendelian randomization study in the Copenhagen General Population Study. *European heart journal* **34**, 2519-2528, doi:10.1093/eurheartj/ehz081 (2013).
- 19 NIH. *NIH to end funding for Moderate Alcohol and Cardiovascular Health trial*, <<https://www.nih.gov/news-events/news-releases/nih-end-funding-moderate-alcohol-cardiovascular-health-trial>> (2018).
- 20 Smith, G. D. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* **32**, 1-22 (2003).
- 21 Davey Smith, G. & Ebrahim, S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* **330**, 1076-1079, doi:10.1136/bmj.330.7499.1076 (2005).
- 22 Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, doi:10.7554/eLife.34408 (2018).
- 23 Burgess, S. & Thompson, S. G. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *American journal of epidemiology* **181**, 251-260, doi:10.1093/aje/kwu283 (2015).
- 24 Sanderson, E., Davey Smith, G., Windmeijer, F. & Bowden, J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol*, doi:10.1093/ije/dyy262 (2018).

- 25 Gage, S. H., Bowden, J., Smith, G. D. & Munafò, M. R. 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 (2018).
- 26 Vu, K. N. *et al.* Causal Role of Alcohol Consumption in an Improved Lipid Profile: The Atherosclerosis Risk in Communities (ARIC) Study. *PloS one* **11**, e0148765-e0148765, doi:10.1371/journal.pone.0148765 (2016).
- 27 Millwood, I. Y. *et al.* Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *The Lancet* **393**, 1831-1842, doi:[https://doi.org/10.1016/S0140-6736\(18\)31772-0](https://doi.org/10.1016/S0140-6736(18)31772-0) (2019).
- 28 Roerecke, M. & Rehm, J. Chronic heavy drinking and ischaemic heart disease: a systematic review and meta-analysis. *Open Heart* **1**, e000135, doi:10.1136/openhrt-2014-000135 (2014).
- 29 Patten, C. A., Martin, J. E. & Owen, N. Can psychiatric and chemical dependency treatment units be smoke free? *Journal of Substance Abuse Treatment* **13**, 107-118, doi:10.1016/0740-5472(96)00040-2 (1996).
- 30 Touchette, J. C. & Lee, A. M. Assessing alcohol and nicotine co-consumption in mice. *Oncotarget* **8**, 5684-5685, doi:10.18632/oncotarget.14603 (2017).
- 31 Marees, A. T. *et al.* Potential influence of socioeconomic status on genetic correlations between alcohol consumption measures and mental health. *Psychol Med*, 1-15, doi:10.1017/s0033291719000357 (2019).
- 32 Karlsson Linnér, R. *et al.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nature Genetics* **51**, 245-257, doi:10.1038/s41588-018-0309-3 (2019).
- 33 Hart, C. L., Davey Smith, G., Gruer, L. & Watt, G. C. The combined effect of smoking tobacco and drinking alcohol on cause-specific mortality: a 30 year cohort study. *BMC public health* **10**, 789, doi:10.1186/1471-2458-10-789 (2010).
- 34 Voight, B. F. *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* **380**, 572-580, doi:10.1016/s0140-6736(12)60312-2 (2012).
- 35 Kosmas, C. E., DeJesus, E., Rosario, D. & Vittorio, T. J. CETP Inhibition: Past Failures and Future Hopes. *Clin Med Insights Cardiol* **10**, 37-42, doi:10.4137/CMC.S32667 (2016).
- 36 Holmes, M. V. *et al.* Mendelian randomization of blood lipids for coronary heart disease. *European heart journal* **36**, 539-550, doi:10.1093/eurheartj/eh571 (2015).
- 37 Barefoot, J. C. *et al.* Alcoholic beverage preference, diet, and health habits in the UNC Alumni Heart Study. *The American journal of clinical nutrition* **76**, 466-472, doi:10.1093/ajcn/76.2.466 (2002).
- 38 Carter, A. R. *et al.* Understanding the consequences of education inequality on cardiovascular disease: mendelian randomisation study. *BMJ* **365**, 11855, doi:10.1136/bmj.11855 (2019).

- 39 Leong, D. P. *et al.* Patterns of alcohol consumption and myocardial infarction risk: observations from 52 countries in the INTERHEART case-control study. *Circulation* **130**, 390-398, doi:10.1161/circulationaha.113.007627 (2014).
- 40 Johnson, C. O. *et al.* Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology* **18**, 439-458, doi:[https://doi.org/10.1016/S1474-4422\(19\)30034-1](https://doi.org/10.1016/S1474-4422(19)30034-1) (2019).
- 41 Babor, T. F. Evidence-based alcohol policy: a call to action. *Addiction* **98**, 1341-1341, doi:10.1046/j.1360-0443.2003.00504.x (2003).
- 42 Prescott, E., Hippe, M., Schnohr, P., Hein, H. O. & Vestbo, J. Smoking and risk of myocardial infarction in women and men: longitudinal population study. *BMJ* **316**, 1043-1047, doi:10.1136/bmj.316.7137.1043 (1998).
- 43 Filion, K. B. & Luepker, R. V. Cigarette smoking and cardiovascular disease: lessons from framingham. *Global heart* **8**, 35-41, doi:10.1016/j.ghheart.2012.12.005 (2013).
- 44 Critchley, J. A. & Capewell, S. Mortality risk reduction associated with smoking cessation in patients with coronary heart disease: a systematic review. *Jama* **290**, 86-97, doi:10.1001/jama.290.1.86 (2003).
- 45 Hackshaw, A., Morris, J. K., Boniface, S., Tang, J. L. & Milenkovic, D. 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 (2018).
- 46 Rosoff, D. B. *et al.* Association of High-Intensity Binge Drinking With Lipid and Liver Function Enzyme Levels]. *JAMA Network Open* **2**, e195844-e195844, doi:10.1001/jamanetworkopen.2019.5844 (2019).
- 47 Agarwala, A. P. *et al.* High-Density Lipoprotein (HDL) Phospholipid Content and Cholesterol Efflux Capacity Are Reduced in Patients With Very High HDL Cholesterol and Coronary Disease. *Arteriosclerosis, thrombosis, and vascular biology* **35**, 1515-1519, doi:10.1161/atvbaha.115.305504 (2015).
- 48 Rohatgi, A. *et al.* HDL Cholesterol Efflux Capacity and Incident Cardiovascular Events. *New England Journal of Medicine* **371**, 2383-2393, doi:10.1056/NEJMoa1409065 (2014).
- 49 Klop, B., Elte, J. W. F. & Cabezas, M. C. Dyslipidemia in obesity: mechanisms and potential targets. *Nutrients* **5**, 1218-1240, doi:10.3390/nu5041218 (2013).
- 50 Volcik, K. A., Ballantyne, C. M., Fuchs, F. D., Sharrett, A. R. & Boerwinkle, E. Relationship of alcohol consumption and type of alcoholic beverage consumed with plasma lipid levels: differences between Whites and African Americans of the ARIC study. *Annals of epidemiology* **18**, 101-107, doi:10.1016/j.annepidem.2007.07.103 (2008).
- 51 Asvold, B. O. *et al.* Causal associations of tobacco smoking with cardiovascular risk factors: a Mendelian randomization analysis of the HUNT Study in Norway. *Int J Epidemiol* **43**, 1458-1470, doi:10.1093/ije/dyu113 (2014).
- 52 He, B. M., Zhao, S. P. & Peng, Z. Y. Effects of cigarette smoking on HDL quantity and function: implications for atherosclerosis. *Journal of cellular biochemistry* **114**, 2431-2436, doi:10.1002/jcb.24581 (2013).

- 53 Gage, S. H. *et al.* Assessing causality in associations between cannabis use and schizophrenia risk: a two-sample Mendelian randomization study. *Psychol Med* **47**, 971-980, doi:10.1017/S0033291716003172 (2017).
- 54 Rosoff, D. B. *et al.* Educational attainment impacts drinking behaviors and risk for alcohol dependence: results from a two-sample Mendelian randomization study with ~780 000 participants. *bioRxiv*, 557595, doi:10.1101/557595 (2019).
- 55 Fawehinmi, T. O., Ilomäki, J., Voutilainen, S. & Kauhanen, J. Alcohol consumption and dietary patterns: the FinDrink study. *PloS one* **7**, e38607-e38607, doi:10.1371/journal.pone.0038607 (2012).
- 56 Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *American Journal of Epidemiology* **186**, 1026-1034, doi:10.1093/aje/kwx246 (2017).
- 57 Ruidavets, J. B. *et al.* Patterns of alcohol consumption and ischaemic heart disease in culturally divergent countries: the Prospective Epidemiological Study of Myocardial Infarction (PRIME). *BMJ* **341**, c6077, doi:10.1136/bmj.c6077 (2010).
- 58 Wray, N. R. *et al.* Pitfalls of predicting complex traits from SNPs. *Nature reviews. Genetics* **14**, 507-515, doi:10.1038/nrg3457 (2013).
- 59 Burgess, S., Davies, N. M. & Thompson, S. G. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* **40**, 597-608, doi:10.1002/gepi.21998 (2016).
- 60 Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat Genet* **45**, 1274-1283, doi:10.1038/ng.2797 (2013).
- 61 Nikpay, M. *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* **47**, 1121-1130, doi:10.1038/ng.3396 (2015).
- 62 Malik, R. *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nature Genetics* **50**, 524-537, doi:10.1038/s41588-018-0058-3 (2018).
- 63 Malik, R. *et al.* Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. *Neurology* **86**, 1217-1226, doi:10.1212/wnl.0000000000002528 (2016).
- 64 Kettunen, J. *et al.* Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nature Communications* **7**, 11122, doi:10.1038/ncomms11122 <https://www.nature.com/articles/ncomms11122#supplementary-information> (2016).
- 65 Yavorska, O. O. & Burgess, S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol* **46**, 1734-1739, doi:10.1093/ije/dyx034 (2017).
- 66 Rees, J. M. B., Wood, A. M. & Burgess, S. 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 (2017).
- 67 Bowden, J. *et al.* A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med* **36**, 1783-1802, doi:10.1002/sim.7221 (2017).

- 68 Bowden, J. *et al.* Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. doi:10.1093/ije/dyy258 (2018).
- 69 Verbanck, M., Chen, C. Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases (vol 50, 693, 2018). *Nature Genetics* **50**, 1196-1196, doi:10.1038/s41588-018-0164-2 (2018).
- 70 Hemani, G., Tilling, K. & Smith, G. D. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *Plos Genet* **13**, doi:ARTN e1007081
10.1371/journal.pgen.1007081 (2017).

TABLES & FIGURES

Table 1. Phenotypes sources

Fig 1. Study overview

Fig 2. Multivariable Mendelian randomization results of alcohol and tobacco use on lipid cardiovascular disease risk factors and disease events

Fig 3. Multivariable inverse variance weighted Mendelian randomization (MVIMR IVW) results for lipoprotein and lipid changes associated with alcohol and tobacco use

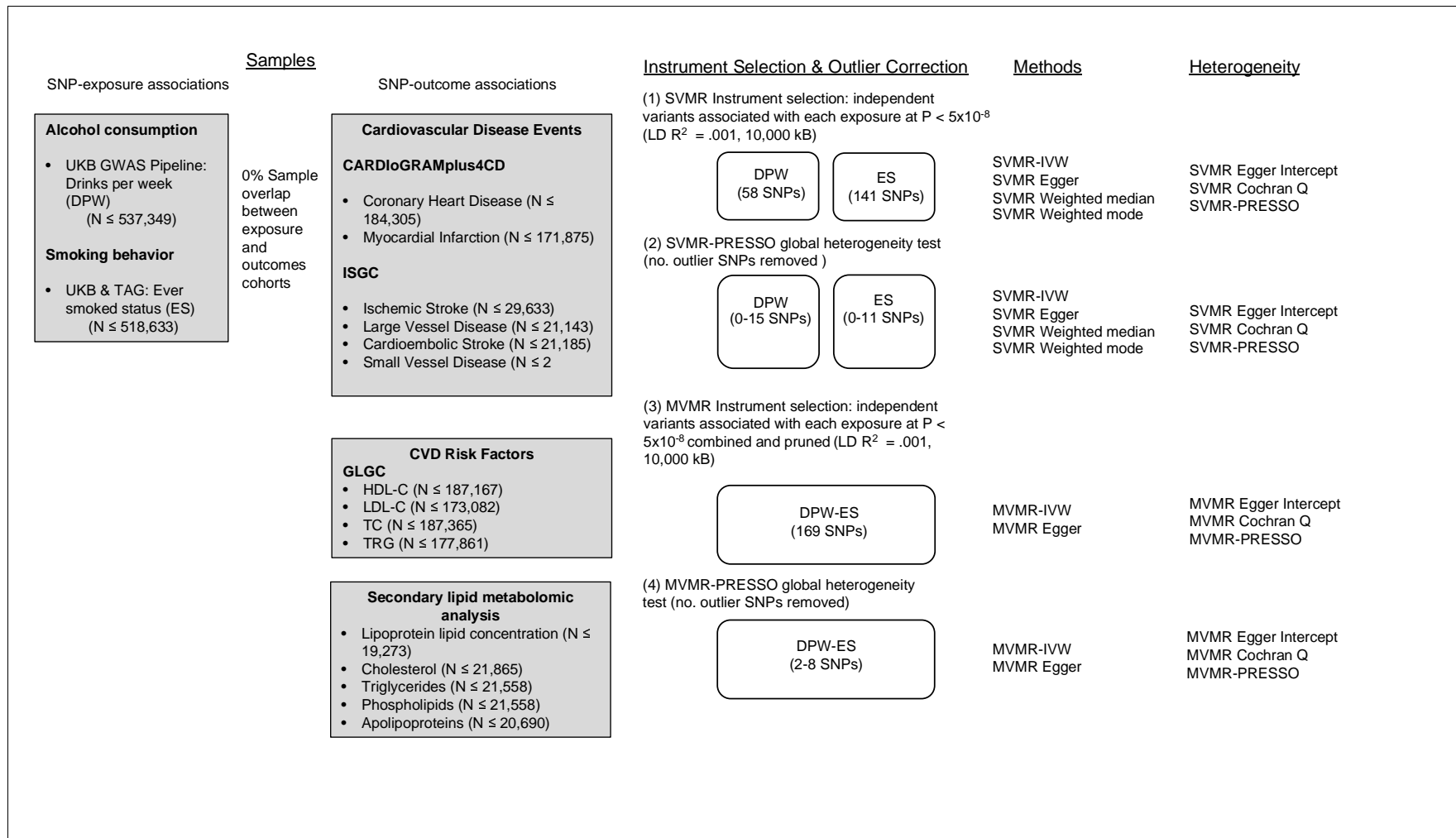


Fig 1. Study overview. Abbreviations: CVD; cardiovascular disease; Dx., diagnosis; GLGC: Global Lipids Genetics Consortium; SSGAC: Social Science Genetic Association Consortium; UKB: UK Biobank; MRC-IEU: Medical Research Center-Integrative Epidemiology Center (UK Bristol);

CARDIoGRAMplus4CD: Coronary ARtery Disease Genome wide Replication and Meta-analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics; IVW, Inverse Variance Weighted MR; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; TC: total cholesterol; TRG: triglycerides; SNP: single nucleotide polymorphism. SVMR; single variable Mendelian randomization; MVMR; multivariable Mendelian randomization.

Table 1. Phenotype sources

Phenotype	Source	First Author (Year)	Sample Size	N Cases	N Controls	Population	Outcome Units
EXPOSURES:							
Alcohol consumption - Drinks Per Week (DPW)	SSGAC (UKB)	Karlsson Linnér (2019)	414343	NA	NA	European	Continuous
Ever Smoker Status (ES)	SSGAC (UKB)	Karlsson Linnér (2019)	518633	NA	NA	European	Binary
PRIMARY OUTCOMES:							
Cardiovascular Disease Risk Factors:							
HDL Cholesterol	GLGC	Willer (2013)	187167	NA	NA	Mixed	SD (mg/dL)
LDL Cholesterol	GLGC	Willer (2013)	173082	NA	NA	Mixed	SD (mg/dL)
Total Cholesterol	GLGC	Willer (2013)	187365	NA	NA	Mixed	SD (mg/dL)
Triglycerides	GLGC	Willer (2013)	177861	NA	NA	Mixed	SD (mg/dL)
Cardiovascular Disease:							
Coronary Heart Disease	CARDIoGRAMplus4CD	Nikpay (2015)	184305	60801	123504	Mixed	Binary
Myocardial Infarction	CARDIoGRAMplus4CD	Nikpay (2015)	171875	43676	128199	Mixed	Binary
Ischemic Stroke	ISGC	Malik (2016)	21185	1859	19326	European	Binary
Cardioembolic Stroke	ISGC	Malik (2016)	21143	1817	19326	European	Binary
Large Vessel Disease	ISGC	Malik (2016)	20675	1349	19326	European	Binary
Small Vessel Disease	ISGC	Malik (2016)	29633	10307	19326	European	Binary
SECONDARY OUTCOMES:							
Circulating Metabolites	NA	Kettunen (2016)	24925	NA	NA	European	SD (mmol/L)

Abbreviations: GLGC: Global Lipids Genetics Consortium; SSGAC: Social Science Genetic Association Consortium; UKB: UK Biobank; MRC-IEU: Medical Research Center-Integrative Epidemiology Center (UK Bristol); CARDIoGRAMplus4CD: Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics; SD: standard deviation (units); mg/dL: milligram per decaliter; mmol/L: millimole per liter; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; SNP: single nucleotide polymorphism; CI: confidence interval.

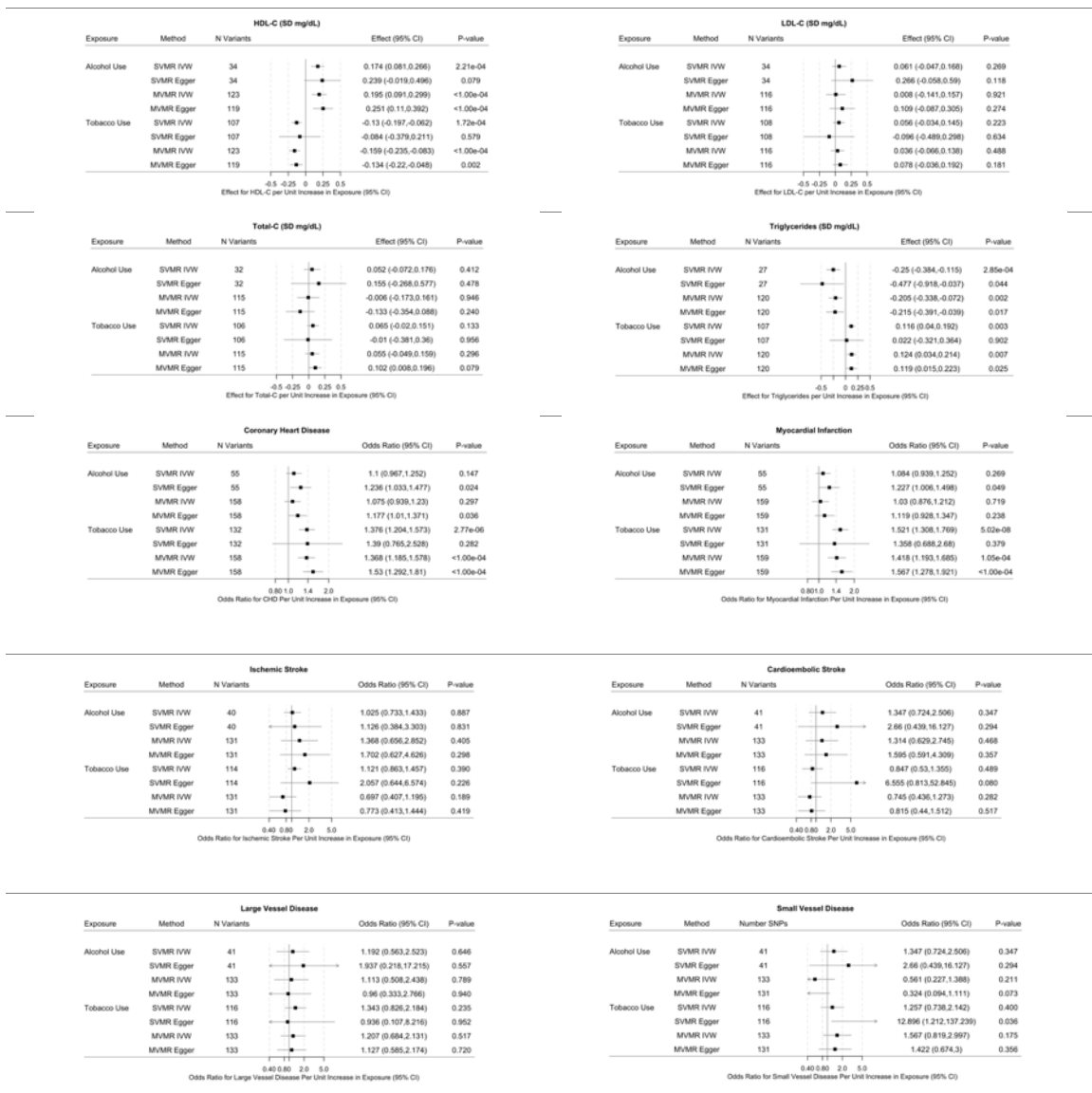


Fig 2. Multivariable Mendelian randomization results of alcohol and tobacco use on lipid cardiovascular disease risk factors and disease events. Potentially pleiotropic outlying SNPs removed. SVMR estimates total effect of a change in each of alcohol and tobacco use in 10 cardiovascular disease risk and event outcomes. MVMR estimates direct effect of alcohol and tobacco use conditional on tobacco and alcohol use respectively. MVMR inverse weighted variance methods shown with complementary MVMR Egger methods for SVMR extended to correct for both measured and unmeasured pleiotropy in MVMR. Full MR results presented in Tables 4, 5 and 6 in Supplement. Abbreviations: SVMR: single variable Mendelian randomization; MVMR: multivariable Mendelian randomization; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; SNP: single nucleotide polymorphism; CI: confidence interval; CVD: cardiovascular disease; IVW: inverse variance weighted MR.

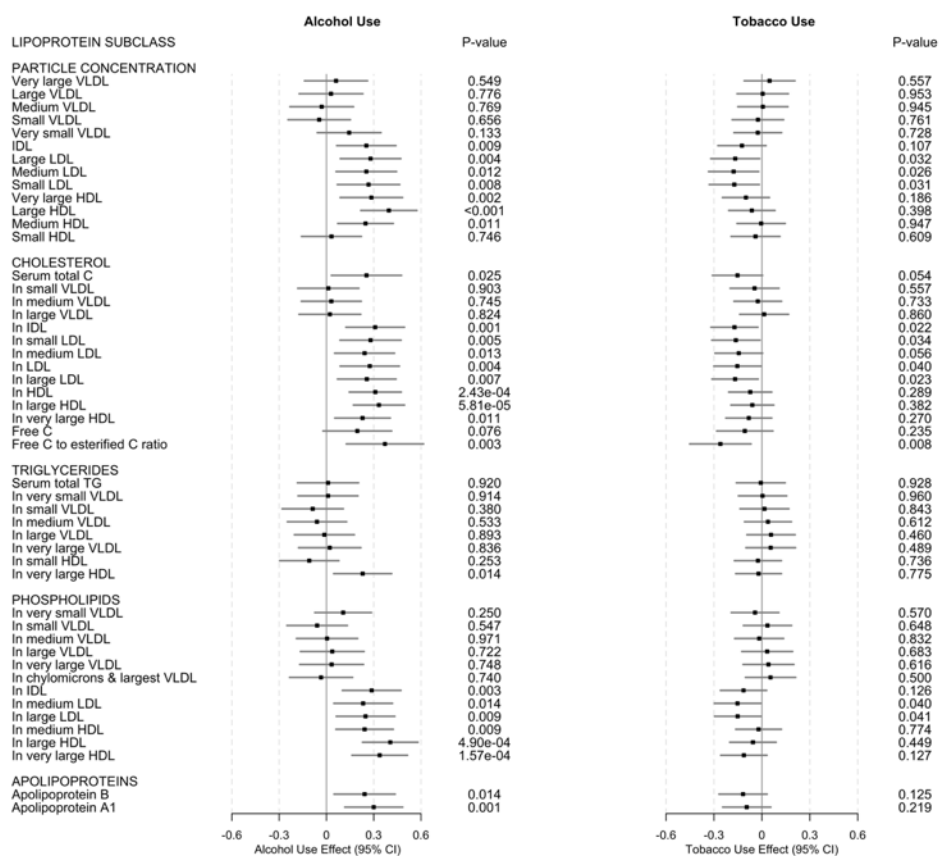


Fig 3. Multivariable inverse variance weighted Mendelian randomization (MVMR IVW) results for lipoprotein and lipid changes associated with alcohol and tobacco use. Results are shown as effects in SD-scaled concentration units per unit change in exposure alcohol or tobacco use. Full IVW MR results in Table 7 in the Supplement. Abbreviations: C: cholesterol; CI confidence interval; HDL: high-density lipoprotein; IDL: IDL: intermediate-density lipoprotein; LDL: low-density lipoprotein; PL: phospholipids; VLDL: very-low-density lipoprotein; TRG: triglycerides.

