Multivariable Mendelian randomization to disentangle the alcohol harm paradox.

Gemma Sawyer¹*, Hannah Sallis¹,², Marcus Munafò¹,³, Liam Mahedy¹, Jasmine Khouja¹,³

¹ MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School,
² Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School,
³ School of Psychological Sciences, University of Bristol, Bristol, United Kingdom

*Corresponding author

Email: gemma.sawyer@bristol.ac.uk

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

The alcohol harm paradox, whereby low socioeconomic position (SEP) groups experience greater alcohol-related harms despite reporting lower alcohol consumption, is yet to be fully understood through observational studies because key drivers are correlated and share similar confounding structures. Multivariable Mendelian randomization (MVMR) were conducted to estimate the direct causal effect of number of drinks per week (DPW) and years of schooling (YOS) on multiple health outcomes. Previously published genome-wide association summary (GWAS) statistics for DPW and YOS were utilised, and summary statistics were generated from individual-level data from UK Biobank (N = 462,818) for all health outcomes. Inverse variance weighted analyses demonstrated evidence for direct effects of DPW and YOS on liver diseases, mental and behavioural disorders due to alcohol, and stroke, indicating that increasing alcohol consumption increased the likelihood of outcomes whereas increasing years of education decreased their likelihood. There was also evidence for a direct effect of DPW on depression, anxiety, influenza/pneumonia, and heart disease. In contrast, there was evidence of a total, but not direct, effect of DPW on depression, influenza/pneumonia, epilepsy, and injuries when accounting for YOS. Although caution is required when interpreting these results due to weak instruments for alcohol, these results provide some evidence that the alcohol harm paradox is partially due to the protective effect of additional years of education, resulting in a reduced likelihood of higher SEP groups developing many alcohol-related outcomes. Replication with strong instruments would be necessary to draw causal inferences.
Author Summary

Individuals from lower socioeconomic backgrounds tend to disproportionately experience alcohol-related physical and mental ill health, despite reporting lower alcohol consumption than those from higher socioeconomic backgrounds. This is known as the alcohol harm paradox. One key difficulty in understanding this paradox is due to methodological limitations establishing the relative contribution of multiple complex social behaviours. In this study, we used genetic variants associated with alcohol consumption and years of education to explore their direct effects on multiple health outcomes. The findings indicate that higher alcohol consumption and fewer years of education may potentially increase the likelihood of developing various health conditions, including liver diseases, mental and behavioural disorders due to alcohol, and stroke. This may suggest that the alcohol harm paradox is due to the protective effect of additional years of education amongst those from higher socioeconomic backgrounds, reducing their likelihood of developing the health conditions. However, these findings are exploratory and limited by various methodological issues, suggesting these findings should be interpreted with caution. Replication and further studies are needed.

Introduction

Alcohol-related harms are common and represent a large burden of morbidity and mortality worldwide, contributing to 5.3% of all deaths and 5.1% of global burden of disease [1]. Moreover, such alcohol-related harms tend to be disproportionately experienced by individuals of low socioeconomic position (SEP) despite reporting lower alcohol consumption compared to high SEP individuals, known as the alcohol harm paradox [2–5]. Despite evidence for this paradox, there is limited understanding of its origins.
There have been multiple explanations proposed for the alcohol harm paradox. One explanation is that, whilst low SEP individuals report consuming less alcohol, other patterns of drinking behaviour are more common. For example, low SEP individuals are more likely to be in the extreme ends of drinking consumption, with a greater proportion abstaining or drinking very lightly compared to high SEP groups, but also being more likely to drink excessively or binge drink [6]. The possibility that drinking patterns may account for the alcohol harm paradox has been supported by research demonstrating a greater socioeconomic gradient for acute health outcomes which are wholly attributable to alcohol, compared to chronic conditions which are associated with a wide range of behaviours [7]. An alternative explanation is that individuals of low SEP may be more likely to engage in other unhealthy behaviours, such as tobacco use, which are associated with multiple adverse health outcomes [8]. Observational studies have supported this explanation, providing evidence that individuals of lower SEP are increasingly likely to smoke [3]. Furthermore, adjusting for tobacco smoking reduces the association between low SEP and hospital admissions [9], while causal mediation analysis demonstrated that 18% of the total effect of income on alcohol-related mortality was due to indirect effects via smoking and BMI [10]. Whilst these proposed explanations have received some support from observational studies, they are limited in their ability to draw causal inferences and are yet to fully account for the alcohol harm paradox.

The primary complication with disentangling the alcohol harm paradox is that a combination of highly correlated causes, including alcohol consumption, SEP, and smoking, could be important drivers. Therefore, it is necessary to establish the relative causal contribution of alcohol consumption alongside these other factors on the likelihood of developing multiple...
adverse health outcomes with novel causal inference methods, such as Mendelian randomization (MR). MR relies on Mendel’s law of random assortment, using genetic variants associated with the exposure as instrumental variables to assess the total causal effect of the exposure on the outcome [11]. Multivariable MR (MVMR) builds upon this method, using genetic variants associated with multiple exposures to estimate the direct causal effect of each exposure on the outcome, accounting for the effect of the other exposures [12]. The validity of causal inferences drawn using MVMR relies on three key assumptions [13]: (1) each exposure is robustly predicted by the genetic instruments; (2) there are no confounders of the outcome and any of the genetic instruments; and (3) the genetic instruments only impact the outcome through at least one of the exposures.

The current study will apply MVMR to the alcohol harm paradox to examine the direct causal effect of alcohol consumption on adverse health outcomes whilst accounting for education (as a proxy for SEP) and vice versa. We will first conduct MR to examine the total causal effect of alcohol and education separately on adverse health, followed by MVMR to examine their direct causal effects. Comparison of the MR and MVMR results will provide insights into the relative contribution of alcohol and education. For example, if there appears to be a causal effect of alcohol in both the MR and MVMR, this suggests that alcohol consumption impacts the health outcome even when the impact of education is considered. However, if a causal effect of alcohol in the MR is attenuated in the MVMR, this may indicate that the impact of education was driving the evidence of a causal effect of alcohol in the MR.

**Materials and Methods**
This method was outlined in our pre-registered analysis plan (https://doi.org/10.17605/OSF.IO/5G8W4) and the methods presented here are consistent with this unless otherwise specified.

Exposures

Summary-level data reported by the two Genome-Wide Association Studies (GWAS) were utilised to identify suitable genetic instruments associated with the relevant exposure phenotypes: drinks per week (DPW) and years of schooling (YOS). In contrast with our pre-registration, the main analysis reported here focuses solely on these two exposures and additional supplementary analysis included cigarettes per day (CPD) as a third exposure (see Supplementary Methods for details). We are focusing on the two exposures of DPW and YOS because alcohol and smoking are highly correlated and both suffer from weak instrument bias in the three-exposure MVMR, making it difficult to draw conclusions from such findings. To aid interpretation, we simplified the analysis to a two-exposure model to explore whether SEP plays a causal role in the alcohol harm paradox. Consequently, the results of this study should be considered exploratory.

GWAS & Sequencing Consortium of Alcohol and Nicotine Use (GSCAN)

Liu and colleagues [14] report summary-level statistics from a GWAS meta-analysis of alcohol use based on data from participants of European ancestry from 29 cohorts (N=941,280). Alcohol use was defined as the reported number of drinks (combined across all alcohol types) consumed per week (DPW) amongst active drinkers. Where individual studies recorded categorical bins of drinks per week (e.g., 1-4, 5-9), the midpoint value was taken. Analyses adjusted for age, sex, age and sex interaction, and 10 genetic principal components.
Independent variants were reported to be associated with DPW if they met the standard significance threshold (p < 5x10^{-8}). The associated SNPs explained 2.5% of the variance in DPW. We utilised the publicly available summary statistics from this GWAS, excluding UK Biobank (UKB), to minimise sample overlap, and 23andMe, due to data sharing restrictions (N=226,223). Smoking (cigarettes per day) was also measured in this GWAS, as described in the Supplementary Material (p. 1).

Social Science Genetic Association Consortium (SSGAC)

Okbay and colleagues [15] report summary-level statistics from a GWAS meta-analysis of education, defined as years of schooling (YOS), based on data from participants of European ancestry from 65 cohorts (N=293,723). Reported number of years of schooling was utilised where this was available. Where individual studies reported qualifications, the International Standard Classification of Education (ISCED; 1997) [16] was used to map each major qualification that could be attained in each country into one of seven categories: pre-primary education, primary education, lower secondary education, upper secondary education, post-secondary non-tertiary education, first-stage tertiary education, and second-stage tertiary education. This was subsequently utilised to impute a YOS value for each category. Analyses adjusted for age, sex, age and sex interaction, ten genetic principal components, and cohort-specific controls. Variants were reported to be associated with YOS if they met the genome-wide significance threshold (p <5x10^{-8}) and were independent.

Health outcomes

Individual-level data, including genetic and outcome data for all individual participants, was obtained from UKB to generate summary statistics for multiple health outcomes.
UK Biobank

UKB is a population-based cohort, which recruited 500,000 people aged between 40-69 years in 2006-2010. With their consent, they provided detailed information about their lifestyle, physical measures and had blood, urine and saliva samples collected and stored for future analysis. Using this individual-level data, we generated summary statistics by regressing each available SNP on each health outcome through the MRC-IEU UKB GWAS pipeline [17] using a BOLT-LMM (linear mixed model) to account for relatedness and population stratification. Furthermore, we additionally adjusted for age, sex, and genotyping chip array. For the main two-exposure MVMR analysis, this was conducted for the overall UKB sample; however, it was also conducted in four smaller samples stratified by smoking status for comparison with the supplementary three-exposure MVMR: ever smokers (defined as those reporting smoking 2 or more cigarettes), never smokers (defined as those reporting smoking 1 or less cigarettes), current smokers, and former smokers (Figure 1).
Figure 1. Flowchart demonstrating the source of exposure and outcome data and the inclusion and exclusion of genetic variants (single nucleotide polymorphisms; SNPs) in this analysis. Details of the three-exposure MVMR (grey dashed lines) are in the Supplementary Materials.
Our pre-registered analysis specified 15 outcomes wholly or partially attributable to alcohol, based upon ICD-10 summary diagnoses classification codes (UKB field ID 41270; see Supplementary Table 1). However, once all the necessary exclusions had been made, four outcomes (alcoholic cardiomyopathy, alcoholic myopathy, alcoholic polyneuropathy, and alcoholic poisoning) had a case number below 50 or less than 0.1% in at least one of the stratified datasets and were therefore removed from the analysis. The remaining outcomes included alcoholic liver disease, other liver diseases, stroke, ischaemic heart disease (IHD), influenza/pneumonia, viral hepatitis, epilepsy, injuries, mental and behavioural disorders due to alcohol, depression, and anxiety.

Analysis

Analyses were conducted in R version 4.1.2 [18], using the Mendelian Randomization[19], TwoSampleMR [20], MVMR [21] and MR-PRESSO [22] R packages.

Selection of genetic variants

Single nucleotide polymorphisms (SNPs) were selected for inclusion in the analysis if they were reported to be associated with the exposure phenotype of interest (alcohol use or education) and (conditionally) independent of any other SNP associations at the genome-wide significance level (p < 5×10^{-8}) in the relevant GWAS. These SNPs also needed to be available in UKB, resulting in 96 DPW and 74 YOS SNPs in the univariable MR. For the MVMR analyses, where a SNP from GSCAN (DPW) was not available in SSGAC (YOS), we selected proxy SNPs with a minimum linkage disequilibrium (LD) R² of 0.8. Moreover, an additional clumping stage was conducted in the MVMR to ensure all exposure SNPs were independent.
(LD $R^2 < 0.1$, clumping window > 500 kb), resulting in 71 DPW and 62 YOS SNPs in the MVMR (Figure 1; see Supplementary Table 2 for details).

We assessed instrument strength using a conditional F-statistic and instrument validity using Cochran’s Q-statistic. In line with conventional guidelines, the F-statistic should be greater than 10 and the Q-statistic should be less than the number of SNPs [23].

Univariable MR

We estimated the total causal effects of alcohol use and education on all 11 health outcomes using two-sample, summary level MR. The main analysis was conducted using inverse variance weighted (IVW) regression [24], however; we also conducted sensitivity analyses using three complementary MR methods MR-Egger [25], weighted median estimation [26], and weighted mode estimation [27]. These methods hold different assumptions regarding pleiotropy and therefore consistent results across the methods provide stronger evidence that the observed effects are causal.

Multivariable MR

In MVMR analyses, we estimated the direct causal effect of alcohol use, education, and tobacco use on all health outcomes using two sample, summary level MVMR with two complimentary methods – MVMR-IVW and MVMR-Egger. Whilst we intended to conduct MVMR-PRESSO in our pre-registered analysis plan, we had to remove this method due to issues with the R package.
Results

The MR-IVW, MVMR-IVW, MR-Egger and MVMR-Egger results exploring the total and direct effects of alcohol and education are displayed in Figure 2. The complete results of the MVMR analyses exploring the direct effects of (1) alcohol and education (n = 133 SNPs) and (2) alcohol, education, and smoking (n = 171 SNPs) on health outcomes are displayed in Supplementary Tables 3-13. The results are presented in odds ratios (ORs) per standard deviation (SD) increase in either drinks per week or years of schooling.

Instrument strength and heterogeneity

Instrument strength was calculated using the conditional F-statistic. In the MR, the F-statistics suggest that the SNPs included in the analysis are strong instruments for alcohol (F = 14.87) and education (F = 35.32). The conditional F-statistics in the two exposure MVMR indicate that the included SNPs are strongly associated with education whilst accounting for the effects of alcohol (F = 17.42), but not strongly associated with alcohol while accounting for education (F = 6.32).

The Cochran’s Q statistics were greater than the number of SNPs included in most of the models, indicating heterogeneity. Therefore, the results also include MVMR-Egger, which provides estimates that are robust to directional pleiotropy under the assumption that pleiotropy is uncorrelated with the strength of the association between the SNP and exposure. However, MVMR-Egger analyses have reduced statistical power compared with IVW analyses and therefore we must consider the results from both analysis when drawing conclusions. Where Egger results indicate a similar effect size but with larger confidence intervals which cross the null, we interpret this as evidence in support of the IVW analysis.
Figure 2. Forest plots displaying the effect of alcohol (drinks per week) and education (years of schooling) on eight health outcomes using Mendelian randomization (MR) and multivariable MR (MVMR) analysis: inverse variance weighted (IVW) and Egger regression.
Alcoholic liver disease

There was evidence for a total (OR = 36.57, 95% CI 16.46 to 81.24) and direct (OR = 50.19, 95% CI 19.35 to 130.21) effect of alcohol on alcoholic liver disease when accounting for education. There was also evidence that education decreased the odds of alcoholic liver diseases in the univariable (OR = 0.44, 95% CI 0.23 to 0.86) and multivariable (OR = 0.27, 95% CI 0.14 to 0.53) analysis when accounting for alcohol (Figure 2; Supplementary Table 3). Whilst the MR-Egger and MVMR-Egger results were consistent with the IVW results for alcohol, for education, the effect estimates were similar but the wider confidence intervals overlapped the null.

Mental and behavioural disorders due to alcohol

There was evidence for a total (OR = 15.88, 95% CI 10.08 to 25.01) and direct (OR = 12.89, 95% CI 7.46 to 22.27) effect of alcohol on mental and behavioural disorders due to alcohol, as well as a total (OR = 0.60, 95% CI 0.44 to 0.83) and direct (OR = 0.51, 95% CI 0.35 to 0.75) effect of education on mental and behavioural disorders due to alcohol (Figure 2; Supplementary Table 6). The MR- and MVMR-Egger findings were consistent with the IVW findings for alcohol and showed similar effect estimates for education, although the confidence intervals crossed the null.

Other liver diseases

There was evidence for a total (OR = 1.80, 95% CI 1.21 to 2.70) and direct (OR = 1.82, 95% CI 1.12 to 2.94) effect of alcohol on other liver diseases, as well as evidence for a total (OR = 0.47, 95% CI 0.35 to 0.62) and direct (OR = 0.42, 95% CI 0.30 to 0.58) effect of education on other liver diseases (Figure 2; Supplementary Table 4). The MR- and MVMR-Egger results
calculated similar effect estimates, although the wider confidence intervals crossed the null for alcohol and the total effect of education.

Stroke

There was evidence for a total (OR = 1.94, 95% CI 1.46 to 2.58) and direct (OR = 1.94, 95% CI 1.30 to 2.89) effect of alcohol on stroke, when accounting for education. Also, whilst there was no clear evidence for a total effect of education on stroke, there was some evidence of a direct effect in the MVMR when accounting for alcohol (OR = 0.73, 95% CI 0.55 to 0.97) (Figure 2; Supplementary Table 5). The MR- and MVMR-Egger findings were similar to the IVW findings in terms of the effect estimates. However, the confidence intervals were wider and therefore crossed the null for education and the direct effect of alcohol.

Injuries

There was evidence for a total (OR = 1.47, 95% CI 1.24 to 1.73) and direct (1.29, 95% CI 1.03 to 1.60) effect of alcohol on injuries. However, there was no evidence for a total (OR = 0.98, 95% CI 0.85 to 1.14) or direct (OR = 0.87, 95% CI 0.75 to 1.02) effect of education (Figure 2; Supplementary Table 7). The MR- and MVMR-Egger results supported the IVW results for education and the total effect of alcohol; however, the confidence intervals crossed the null for the direct effect of alcohol.

Depression

In the univariable MR, there was evidence for a total effect of alcohol on depression (OR = 1.33, 95% CI 1.09 to 1.64); however, there was no evidence for a direct effect when accounting for education (OR = 1.11, 95% CI 0.84 to 1.48). There was evidence that education
decreased the odds of depression in the univariable (OR = 0.59, 95% CI 0.48 to 0.73) and multivariable analyses when accounting for alcohol (OR = 0.56, 95% CI 0.45 to 0.68) (Figure 2; Supplementary Table 8). The MR- and MVMR-Egger results supported the IVW findings of a total, but not direct, effect of alcohol (although the wider confidence intervals crossed the null), as well as a direct effect of education. The MR-Egger did not however support IVW evidence for a total effect of education.

Epilepsy

Whilst there was evidence for a total effect of alcohol on epilepsy (OR = 1.91, 95% CI 1.33 to 2.75), there was no evidence for a direct effect when accounting for education (OR = 1.17, 95% CI 0.72 to 1.89). Moreover, there was no evidence for a total (OR = 0.83, 95% CI 0.60 to 1.15) or direct (OR = 0.82, 95% CI 0.58 to 1.14) effect of education on epilepsy (Figure 2; Supplementary Table 9). The MR-Egger results supported a total effect of alcohol and also found some evidence for a direct effect, which was not indicated in the IVW findings. The MR- and MVMR-Egger effect estimates for education were in the opposite direction to the IVW results, although all confidence intervals crossed the null.

Influenza and pneumonia

There was evidence for a total (OR = 1.32, 95% CI 1.03 to 1.68) but not a direct effect of alcohol on influenza/pneumonia when accounting for education (OR = 1.28, 95% CI 0.94 to 1.73). There was evidence that education decreased the odds of influenza/pneumonia in the univariable (OR = 0.63, 95% CI 0.53 to 0.76) and multivariable (OR = 0.60, 95% CI 0.48 to 0.74) analysis (Figure 2; Supplementary Table 10). The MR- and MVMR-Egger results were primarily consistent with the IVW results, supporting a total and direct effect of education and no direct
effect of alcohol. Whilst the effect estimate for the total effect of alcohol was similar, the MR-
Egger confidence interval crossed the null.

Other outcomes

There was no evidence for a total or direct effect of alcohol on IHD (Supplementary Table 11),
viral hepatitis (Supplementary Table 12), or anxiety (Supplementary Table 13). There was
evidence for a total (OR = 0.61, 95% CI 0.53 to 0.70) and direct (OR = 0.62, 95% CI 0.52 to 0.74)
effect of education on IHD when accounting for alcohol. There was also evidence for a total
(OR = 0.67, 95% CI 0.52 to 0.86) and direct (OR = 0.62, 95% CI 0.48 to 0.80) effect of education
on anxiety. There was no evidence for a total or direct effect of education on viral hepatitis.

Overall, the MR- and MVMR-Egger results were mostly similar to the IVW results for IHD,
anxiety, and viral hepatitis. The only differences were that the MR-Egger did not support a
total effect of education on IHD or anxiety.
Discussion

The current study has found evidence for direct effects of alcohol and education on alcoholic and other liver diseases, mental and behavioural disorders due to alcohol, and stroke, indicating that alcohol use increases the likelihood of such outcomes whereas years of education decreases their likelihood. There was also evidence of a direct effect of education on depression, anxiety, influenza/pneumonia, and IHD. In contrast, there was evidence of a total, but not direct effect of alcohol on depression, influenza/pneumonia, epilepsy, and injuries when accounting for education. These results are consistent with previous findings from observational studies suggesting that both alcohol consumption and SEP may impact the likelihood of developing numerous health conditions and that the alcohol harm paradox is due to these multiplicative risk factors [3,4]. However, when interpreting these results, it is important to be cautious and consider the strength of instruments. The conditional F-statistics indicate that the instruments were not strongly associated with alcohol in the MVMR, despite strong associations in the MR. Therefore, we cannot be as confident or draw causal inferences from our estimates of direct effects; instead, our findings may be indicative of possible causal effects which would need to be replicated with strong instruments to be confident. In addition, we can only consider these results as exploratory as they were not all specified in our pre-registered analysis plan.

For multiple outcomes, there was little difference between the total and direct effects of both alcohol and education, including alcoholic liver disease, other liver diseases, and mental and behavioural disorders due to alcohol. These results suggest that alcohol use increases the likelihood of such outcomes when accounting for education, whereas years of education decreases their likelihood when accounting for alcohol use. Therefore, for these outcomes, it
is possible that greater alcohol consumption and less education both causally contribute to their likelihood and, as such, the alcohol harm paradox may be due to a combination of these risk factors in low SEP groups, whereas high SEP groups are somewhat protected by additional years of education. Moreover, there was evidence for a total and direct effect of alcohol on stroke but only evidence for a direct, but not a total, effect of education. Again, these results indicate that alcohol use increased the likelihood of stroke, even when accounting for education, and education decreases the likelihood of stroke. However, it may also suggest that failing to account for alcohol in the MR masked the effect of education on stroke.

For depression and influenza/pneumonia, there was little difference between the total and direct effects of education; however, the evidence for a total effect of alcohol was attenuated when accounting for education. This suggests that years of education decreases the likelihood of developing depression and influenza/pneumonia, whereas alcohol use does not directly influence the likelihood of such outcomes when education is considered alongside. Similarly, for injuries and epilepsy, there was evidence for a total effect of alcohol which was attenuated when education was considered, perhaps suggesting that alcohol does not directly influence these outcomes. However, there was no evidence of a total or direct effect of education on injuries and epilepsy these outcomes, indicating that education also does not influence these outcomes. It is plausible that the attenuation of the total effect of alcohol is the result of reduced instrument strength in the MVMR compared with the MR as opposed to education driving the total effect of alcohol identified in the MR. Replication of our results with stronger instruments is necessary to clarify such findings. The current results also support a total and direct effect of education only on anxiety and IHD, whereby years of education decreased the likelihood of developing these outcomes when accounting for alcohol. There was no evidence
for a total or direct effect of alcohol on either anxiety or IHD, nor a total or direct effect of alcohol or education on viral hepatitis.

Strengths and Limitations

This is the first study, to our knowledge, to apply MVMR, a novel causal inference method, to understand the direct causal effects of alcohol and education on multiple alcohol-related conditions to enhance understanding of the alcohol harm paradox. However, this study also has multiple limitations. As previously mentioned, the conditional F-statistics indicate that alcohol suffers from weak instruments in the MVMR, which may possibly bias our results. Heterogeneity is indicated by the Cochran’s Q statistics and, even though we applied MR-Egger to address this, heterogeneity may remain. This indicates that the analyses may also be biased by invalid instruments. Directional pleiotropy is similarly likely to persist in our analyses despite applying pleiotropy-robust methods as indicated by the MR-Egger intercepts, which are not null in most of the models. This means that the genetic instruments may influence the adverse health outcomes within this study via pathways that do not include alcohol or education and therefore may not tell us about the causal effect of alcohol or education on such outcomes. Moreover, DPW is a somewhat limited phenotype to reflect alcohol. This is because many factors influence the effects of alcohol beyond number of drinks, including binge patterns and alcohol types. Similarly, education is one factor that contributes to an individual’s SEP; factors such as income, neighbourhood, and occupation are also important and may show different relationships. Therefore, focusing purely on DPW and YOS may be restricting our understanding. Combined with the exploratory nature of these analyses, these limitations suggest that caution is highly necessary in interpreting the current findings.
Future Research and Implications

Due to limitations within the study regarding weak instruments and possible pleiotropy, the current study should be replicated with newly available GWAS data [28] and thus potentially valid, stronger instruments can be utilised in these analyses. Replication will help confirm whether the results presented in this study reflect the true causal effects of alcohol and education or are due to methodological limitations. If the current findings were supported, this would suggest that the alcohol harm paradox may be due to the protective effect of years of schooling on many alcohol-related conditions resulting in a reduced likelihood of higher SEP groups developing such outcomes. Future research may therefore benefit from exploring the mechanisms through which education impacts such health outcomes to inform preventative policy. Beyond this, it may be useful for future work to explore alternative exposures, such as type of alcohol and income, and outcomes, including cancers to broaden understanding.

Conclusion

This study found evidence for a direct effect of alcohol on alcoholic and other liver diseases, mental and behavioural disorders due to alcohol, and stroke. There was also evidence for a total effect of alcohol on depression, epilepsy, influenza/pneumonia, and injuries, which were attenuated when education was accounted for. Despite weak instrument bias, such results cautiously indicate that many partially alcohol-attributable conditions are causally impacted by years of education, which we use here as a proxy for SEP. Replication with strong instruments is necessary however to confirm these findings and future research may consequently endeavour to explore the causal effect of education on additional partially
alcohol-attributable conditions and identify the mechanism through which further years of education protects against these conditions.

Acknowledgements

This research has been conducted using the UK Biobank Resource under application number 9142. This work uses data provided by patients and collected by the NHS as part of their care and support.

Funding

GS, HS, MM, LM, and JK are all members of the Medical Research Council (MRC) Integrative Epidemiology Unit at the University of Bristol. This publication is the work of the authors and they will serve as guarantors for the contents of this paper. GS is supported by a Wellcome Trust PhD studentship in Molecular, Genetic and Lifecourse Epidemiology (ref: 218495/Z/19/Z). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.
References


