 doi:10.20944/preprints202205.0258.v1

Molecular alterations caused by alcohol consumption: A Mendelian Randomisation study

Felix O'Farrell [1], MSc; Xiyun Jiang [1], MSc; Shahad Aljifri [1], MSc; Raha Pazoki [1,2], MD and PhD

1. Division of Biomedical Sciences, Department of Life Sciences, College of Health and Life Sciences, Brunel University, London, UB8 3PH, United Kingdom
2. MRC Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, St Mary's campus, Norfolk Place, London W2 1PG, United Kingdom

* Correspondence: author: Raha Pazoki, Division of Biomedical Sciences, Department of Life Sciences, College of Health and Life Sciences, Brunel University, London, UB8 3PH, United Kingdom; Email: raha.pazoki@brunel.ac.uk

Abstract: Background: Alcohol consumption is associated with the development of cardiovascular diseases, cancer, and liver disease. The biological mechanisms are still largely unclear. Here, we aimed to use an agnostic approach to identify phenotypes mediating the effect of alcohol on various diseases.

Methods: We performed an agnostic association analysis between alcohol consumption (red, and white wine, beer/cider, fortified wine, and spirits) with over 7,800 phenotypes from the UK biobank comprising 223,728 participants. We performed Mendelian randomisation analysis to infer causality. We additionally performed a Phenome-wide association analysis and a mediation analysis between alcohol consumption as exposure, traits in causal relationship with alcohol consumption as mediators, and various diseases as outcome.

Results: Of 45 traits in association with alcohol consumption, 20 were in causal relationship with alcohol consumption. Gamma glutamyltransferase (GGT; β=9.44; CI,5.94-12.93; P_{fdr}=9.04×10^{-7}), mean sphered cell volume (β=0.189; CI,0.11-0.27; P_{fdr}=1.00×10^{-4}), mean corpuscular volume (β=0.271; CI,0.19-0.35; P_{fdr}=7.09×10^{-10}) and mean corpuscular haemoglobin (β=0.278; CI,0.19-0.36; P_{fdr}=1.60×10^{-6}) showed the strongest causal relationships. We also identified GGT and physical activity as mediators causing liver cirrhosis and alcohol dependence.

Conclusion: Our study provides evidence of causality between alcohol consumption and 20 traits and a mediation effect for physical activity on health consequences of alcohol consumption.

Keywords: Mendelian randomisation; Alcohol Consumption; UK Biobank; Phenome wide association studies; Biomarker

Introduction

Alcohol use is responsible for 5.1% of the global burden of disease (Organization World Health, 2018) and is considered as the main contributor to alcohol liver disease (ALD). Despite a general understanding of the link between alcohol consumption and diseases, the causal associations and mediatory mechanisms are less clear.

Phenome wide association studies (PheWAs) examine correlation between an exposure (a variant or phenotype of interest) with an array of outcomes (the phenome) (Anurag Verma & Marylyn D. Ritchie, 2017).

 doi:10.20944/preprints202205.0258.v1

The power of a PheWAS is determined by the sample size and variety of clinical information present in the database which makes up the phenome (J. C. Denny et al., 2010). The current gold-standard to perform PheWAS are large and comprehensive electronic health record datasets (Cathie Sudlow et al., 2015).

PheWAS establish associations between exposures and outcomes and are unable to assess causal links. Mendelian randomisation (MR) studies investigate causality between instruments and often follow PheWAS to further interrogate a suggested association. MR studies operate upon the fact that most genetic variants are inherited randomly from parents and can be used as a randomisation tool to mimic randomised clinical trials (George Davey Smith & Gibran Hemani, 2014). Owing to the random nature of genetic inheritance, MR studies decrease bias such as reverse causality and confounding (Jie Zheng et al., 2017). Genetic variants which are associated with traits (found from association studies) can be used as exposure instruments to test for causal associations against corresponding outcome instruments (George Davey Smith & Gibran Hemani, 2014).

Identification of causal biomarkers in the pathway between alcohol consumption and alcohol liver disease could facilitate their utility in better identification of individuals at higher risk of developing alcohol-related diseases. These high-risk individuals could then be targeted to receive public health interventions.

Here, we applied a multi-stage design to identify traits in causal pathway between alcohol consumption and alcohol-related diseases. Using a combination of agnostic approaches, PheWAS and MR analyses, we investigated over 7,800 phenotypes in the UK biobank cohort for association and causal link with alcohol consumption and alcohol related diseases.

Methods

We used data from the UK biobank (n=223,728) which is a large biomedical database with genotypic and phenotypic data on a wide range of health-related outcomes for over 500,000 individuals. Participants with aged between 40 to 69 and lived within 30 miles of one of the 22 UK biobank assessment centres were invited to take part. The UK biobank has full ethical approval by the UK NHS National Research Ethics Service (Cathie Sudlow et al., 2015). All participants in this study gave consent for their data to be used (Fidel Alfaro-Almagro et al., 2021; Simon D. Kyle et al., 2017).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2022 doi:10.20944/preprints202205.0258.v1

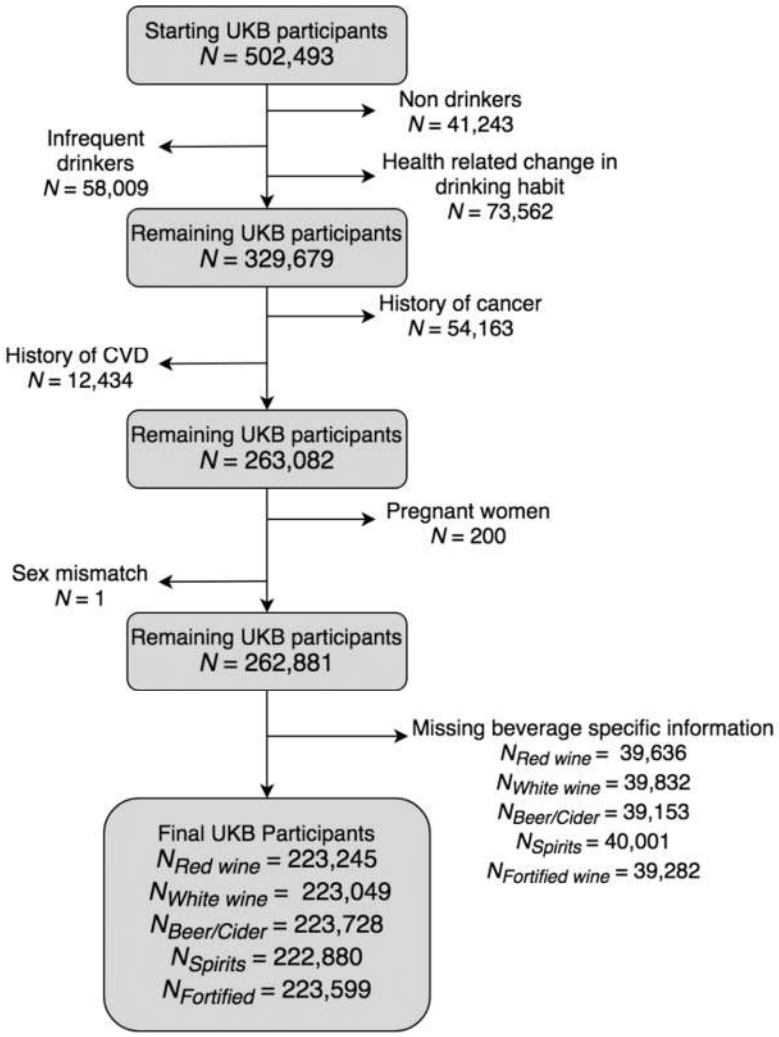

Figure 1. Overview of included and excluded participants.

Participants who stated drinking of at least one of the alcoholic beverages red wine, white wine, beer/cider, spirits, and fortified wine were included in the current analysis (**Figure 1**). Participants who withdrew consent were removed from the dataset (*N*=109), leaving a starting total of *N*=502,493. Individuals who completed the UK biobank touchscreen questionnaire and declared they do not drink (*N*=41,243), only drink on special occasions (*N*=58,009) or changed drinking habits due to health reasons (*N*=73,562) were removed from this analysis. This ensured our analyses focused on individuals who drink alcoholic beverages on a regular basis. Using the same questionnaire, individuals who had serious co-morbidities a baseline such as self-reported cancer (*N*=54,163) and cardiovascular disease (*N*=12,434) were also excluded. Participants who had missing data for the alcohol phenotypes of interest, pregnant women (*N*=200) and individuals who had missing sex data (*N*=1) were excluded. Individuals who passed the exclusion criteria but did not have beverage specific data were removed from this analysis. This slightly varied depending on the beverage type (*N*~44,758). Final beverage specific data used in our analysis include red wine (N= 223,245), white wine or champagne (N= 223,049), beer or cider (N= 223,728), fortified wine (N= 223,599) and spirits (N= 222,880).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2022 doi:10.20944/preprints202205.0258.v1

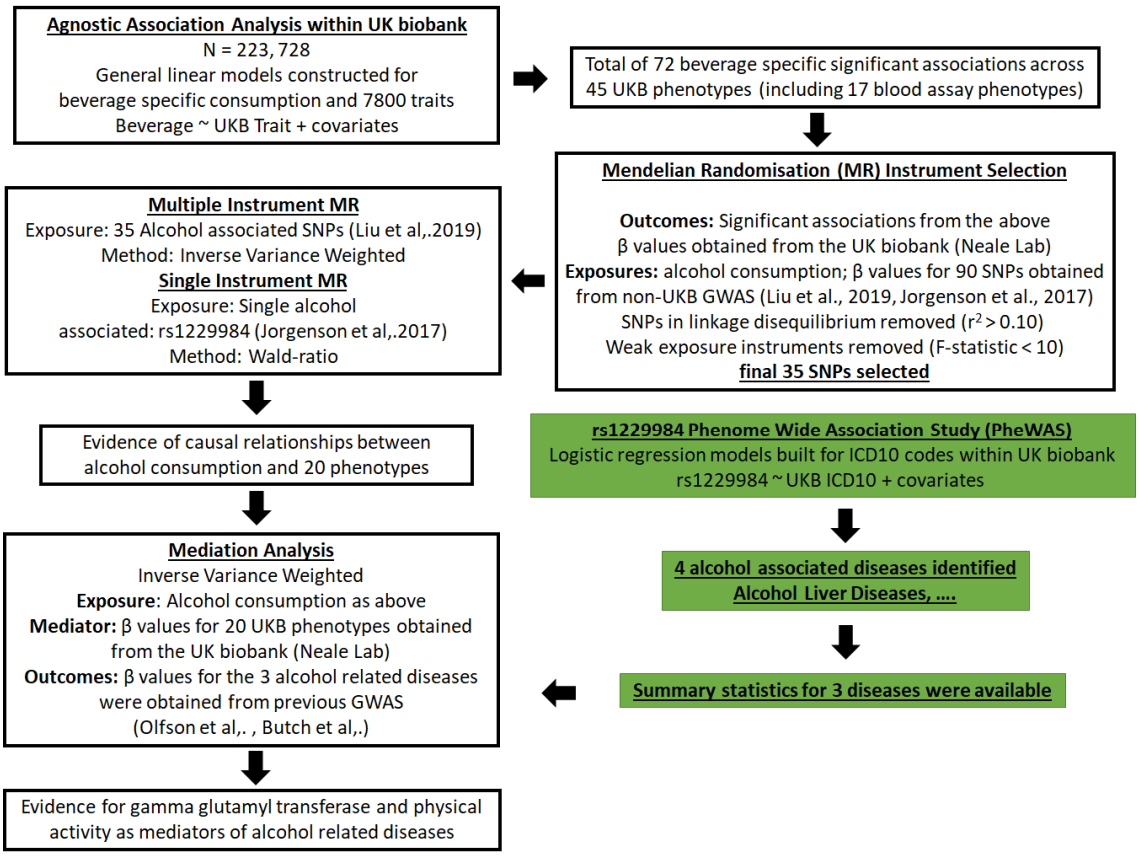

Figure 2. Overview of the study design.

The overview of the study design is shown in **Figure 2**. Data collection was performed centrally by the UK biobank. Between 2006 and 2010, touchscreen questionnaires and in-person interviews were conducted by UK biobank in one of their 22 UK-wide assessment centres. These baseline assessments include lifestyle choices, environmental factors along with personal and family medical history. UK biobank participants also gave blood and urine samples (Cathie Sudlow et al., 2015). All, biological samples were obtained and analysed using the protocol outlined by the UK biobank (T. C. Peakman & P. Elliott, 2008). Diagnosis of diseases were made based on International Classification of Diseases (ICD 10) coding within the UK biobank data (https://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=41203).

Self-reported information describing the weekly frequency of different kinds of alcoholic beverages have been collected in the UK biobank. Participants specified their consumption based on the number of glasses of red wine, white wine/champagne and fortified wine in an average week. Number of pints of beer/cider consumed per week and measures of spirits or liquors consumed per week were collected to assess consumption of beer/cider and spirits.

Beverage specific agnostic association analyses

We performed an initial agnostic analysis in which over 7,800 phenotypes and circulatory biomarkers from the UK biobank were used to investigate association with drinking alcoholic beverages.

 doi:10.20944/preprints202205.0258.v1

We performed five agnostic association analyses (one for each alcoholic beverage) across the UK biobank database. Weekly consumption of red wine, white wine/champagne, beer/cider, spirits and fortified wine acted as individual outcome variables in each agnostic association analysis. Individual general linear models were constructed for each outcome variable with one of the 7,803 UK biobank phenotypes acting as the exposure variable. In our agnostic association approach, we performed a linear regression model adjusted for potential confounders (age, sex, Townsend deprivation index, genetic and ethnic background, smoking status, and diabetes).

To test the significance for each model we calculated an empirical P-value using a 10,000-iteration permutation test (A. C. Davison & D. V. Hinkley, 1997). This involves calculating the probability (empirical *P*-value) that the observed *P*-value from a given model is driven by chance. To calculate the empirical *P*-value, we randomly sampled the outcome variable and calculated the P value for the association of alcoholic beverages consumption with this randomly sampled outcome. This was repeated for 10,000 iterations and the *P*-value obtained from each of these iterations was estimated. The number of permuted *P*-values which were less than the observed (n) +1 was then divided by the number of iterations of the permutation test (s) +1 based on the equation:

$$P_e = \frac{n+1}{s+1}$$

Where P_e is the empirical *P*-value, n is the number of permuted *P*-values less than the observed *P*-value and s is the number of iterations of the permutation test. If the empirical *P*-value was less than 0.05, the model was considered significant (A. C. Davison & D. V. Hinkley, 1997). Analysis of Variance (ANOVA) were conducted on all significant models and the percentage of explained variance in the alcoholic beverage consumption by each UK biobank phenotype was obtained.

Two Sample MR

MR analysis uses genetic variants to explore causal relationships between an exposure and outcome. To infer causality between the effect of alcohol consumption and other traits, we performed two sample MR analysis. In our MR analyses, alcohol consumption was used as the exposure and genetic variants associated with alcohol consumption were used as instrumental variable. All non-alcoholic traits that were significantly associated with alcohol consumption in our initial agnostic association analysis were considered as outcome for the MR analysis. *i.e.*, traits that were significantly associated with consumption of any alcoholic beverages in our agnostic analysis were moved forward to be assessed for causality of the association with alcohol consumption. MR analyses were performed using either multiple alcohol associated instruments (Mengzhen Liu et al., 2019) or the single rs1229984 instrument (E. Jorgenson et al., 2017).

- Instrument selection (Multi-instrument MR)

We obtained genetic data from previously performed genome-wide association analyses (GWAS) for the effect of genetic variants on alcohol consumption (E. Jorgenson et al., 2017; Mengzhen Liu et al., 2019). β value for each of the genetic variant associated with alcohol consumption was

 doi:10.20944/preprints202205.0258.v1

used as an instrumental variable (alcohol consumption was defined as the exposure; see above). Weak instruments are genetic variants that do not capture enough variance in an exposure and introduce bias into a MR analysis (S. Burgess & S. G. Thompson, 2011). To account for weak instrument bias, the β values were used to calculate the phenotype proportion of variance explained (R^2) along with the F-statistic for each instrument (**Supplementary Table 1**). The F-statistic is a measure of the association between the SNP and the exposure. Any instrument scoring an F-statistic of less than 10 was removed from the analysis (Tanya M. Teslovich et al., 2010).

In addition, genetic variants that were in linkage disequilibrium with other variants (indicated by $r^2<0.1$) were identified using the European population data from the 1000 genomes project (Laura Clarke et al., 2012) and were excluded from the instrument list.

Liu and colleagues performed meta-analyses across multiple alcohol consumption GWA studies including the UK biobank. To avoid sample overlap with UK biobank, we selected the β values for alcohol consumption from one of the meta-analyses by Liu et al. that excluded UK biobank data.

To obtain the β value for each genetic variant associated with outcomes (shortlisted traits from agnostic models), we used summary statistics from Neale Lab's UK biobank GWAS studies (http://www.nealelab.is/). Neale lab has performed linear regression GWAS on almost all UK biobank phenotypes, making it a consistent set of summary statistics to use for our MR analyses.

The R package TwoSampleMR was used to harmonise the exposure and outcome effect estimates (Gibran Hemani et al., 2018**).** The Inverse Variance Weighting (IVW) method implemented in the MR-PRESSO package was used to perform the MR analyses and identify outlier genetic variants responsible for horizontal pleiotropy and remove them from the analysis (Marie Verbanck, Chia-Yen Chen, Benjamin Neale, & Ron Do, 2018). A false discovery rate (FDR) of 0.05 was used to adjust the *P*-values from all analyses (To account for multiple testing).

- Single-instrument MR analysis

As a sensitivity analysis for our causal inference, we performed a single-instrument MR analysis in which, the β value for the association of rs1229984 with alcohol consumption was obtained from the study by Jorgenson and colleagues (E. Jorgenson et al., 2017). rs1229984 occurs in *ADH1A* gene as is known to be a functional non-synonymous (Arg48His) genetic variant for alcohol consumption (Eric J. Duell et al., 2011). *ADH1B* gene encodes for an enzyme responsible for oxidising alcohol (Howard J. Edenberg & Jeanette N. McClintick, 2018).The Wald ratio method was used for our single-instrument MR analysis.

Results that consistently showed causal relationship with the same direction of effect across both multiple and single-instrument MR analysis were considered significant.

Phenome-wide association analysis

We performed a Phenome-wide association analysis (PheWAs) on rs1229984 as our exposure within the UK biobank. This was done to identify alcohol-related diseases. We used Logistic regression models within the R package PheWAS. Genotypes for the rs1229984 were extracted from individual level data of the UK biobank using plink

 doi:10.20944/preprints202205.0258.v1

(Shaun Purcell et al., 2007). A list of diagnosed diseases and conditions from hospital episode statistics were available within the UK biobank data in the form of ICD10 codes. We adjusted for the same potential confounders that were included in our agnostic analyses. To account for multiple testing, a Bonferroni correction was implemented.

Mediation analysis

We performed a mediation analysis, to investigate whether the traits identified to be causally linked to alcohol consumption through our MR analysis mediate the effect of alcohol consumption on alcohol -related diseases. Diseases that were linked to rs1229984 within our PheWAS analysis were included in our mediation analysis as outcomes.

Different steps of the mediation analysis included investigation of causality between alcohol consumption and alcohol-related diseases mediated by the UK biobank phenotypes from our agnostic approach. Causal link between alcohol and the UK biobank phenotypes was demonstrated in our agnostic approach. To show the causal link between the UK biobank phenotypes and alcohol-related diseases, we performed another set of MR analysis. Here, we used Neale lab GWAS data to identify and select SNPs associated with mediators (i.e. traits highlighted to be linked to alcohol consumption in our agnostic approach).

We additionally used GWAs performed on alcohol associated disease (L. J. Bierut et al., 2010; S. Buch et al., 2015; Tuomo Kiiskinen et al., 2020) to perform mediation analysis between mediators and alcohol-related diseases. QC thresholds of linkage disequilibrium and instrument strength that were applied to this set of mediation analysis were described in the instrument selection section (see above).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2022 doi:10.20944/preprints202205.0258.v1

Table 1: Baseline characteristics of the population for analysis.

Characteristics	Red wine (n=223,245)	White wine/sparkling white wine (n=223,049)	Beer or cider (n=223,728)	Spirits (n=222,880)	Fortified wine (n=223,599)
Age-yr	55.5 (±8.01)	55.5 (±8.01)	55.5 (±8.01)	55.5 (±8.01)	55.593 (±8.01)
Male sex -no. (%)	108467 (48.59%)	108509 (48.61%)	108579 (48.64%)	108529 (48.61%)	108458 (48.58%)
Lipid treatment-no./total no(%)	24259 (10.87%)	24246 (10.86%)	19778 (8.86%)	19756 (8.85%)	19774 (8.86%)
Diabetes mellitus-no./total no. (%)	5840 (2.62%)	5829 (2.61%)	5847 (2.62%)	5830 (2.62%)	5837 (2.61%)
Body mass index	26.9 (±4.33)	26.9 (±4.33)	26.9 (±4.34)	26.9 (±4.33)	26.9 (±4.34)
MET Score	2642.4 (±2664.35)	2642.8 (±2665.44)	2642.5 (±2665.28)	2642.1 (±2664.77)	2641.4 (±2663.52)
current smoking-no. (%)	23593 (10.57%)	23599 (10.57%)	23659 (10.60%)	23553 (10.55%)	23620 (10.58%)
past smoking-no. (%)	77860 (34.88%)	77884 (34.89%)	77875 (34.88%)	77881 (34.89%)	77875 (34.88%)
never smoking-no. (%)	121130 (54.26%)	121097 (54.24%)	121044 (54.22%)	121148 (54.27%)	121084 (54.24%)
Systolic blood pressure-mean (SD)- mmHg	140.4 (±19.47)	140.4 (±19.48)	140.4 (±19.48)	140.4 (±19.48)	140.4 (±19.49)
Diastolic blood pressure-mean (SD)- mmHg	83.4 (±10.82)	83.4 (±10.81)	83.4 (±10.81)	83.4 (±10.81)	83.4 (±10.82)
Red wine intake- mean (SD)-glass/day	3.9 (±5.68)	3.9 (±5.68)	3.9 (±5.68)	3.93 (±5.68)	3.9 (±5.68)
White wine intake- mean (SD)- glass/day	2.7 (±4.88)	2.7 (±4.88)	2.7 (±4.88)	2.7 (±4.88)	2.7 (±4.88)
Fortified wine intake- mean (SD)- glass/day	0.2 (±1.21)	0.2 (±1.21)	0.2 (±1.22)	0.2 (±1.22)	0.2 (±1.22)
Beer intake- mean (SD)-glass/day	2.9 (±5.59)	2.9 (±5.59)	2.9(±5.62)	2.9 (±5.60)	2.9 (±5.60)
Spirits intake- mean (SD)-glass/day	1.8 (±5.29)	1.8 (±5.29)	1.8 (±5.32)	1.8 (±5.36)	1.8 (±5.30)

 doi:10.20944/preprints202205.0258.v1

Results

Our data for analysis included 223,728 individuals from the UK biobank of which 48% were males (**Table 1**). The average age in the cohort used in our analysis was 55.5 (±8.01) and the average body mass index (BMI) was 26.9 (±4.33). We found 45 traits (**Figure 3**). significantly associated with consumption of at least one alcoholic beverage (beer/cider, white wine, red wine, fortified wine, and spirits).

We observed association between gamma glutamyl transferase (GGT) with increased consumption of beer/cider (β =0.02; s.e =2.66×10^{-4}; P <1.0×10^{-300}; r^2=5.3%) and spirit (β =0.01; s.e =2.9×10^{-4}; P <1.0×10^{-300}; r^2=1.39%). Insulin-like Growth Factor (IGF-1) showed a negative association with beer/cider consumption (β =-0.09; s.e=2.05×10^{-3}; P <1.0×10^{-300}; r^2=0.21%) (**Figure 4**). Individuals showed a significant increase in the levels of apolipoprotein A1 (apo-A1) in their blood assays if they reported a higher weekly consumption of beer or cider (β =2.2; s.e =4.6×10^{-2}; P <1.0×10^{-300}; r^2=0.89%), white wine (β =0.09; s.e =2.1×10^{-3}; P <1.0×10^{-300}; r^2=2.89%) and red wine (β =3.1; s.e =5.2×10^{-2}; P <1.0×10^{-300}; r^2=0.71%). We additionally observed association between cystatin c levels in the blood and red wine consumption (β =-4.0; s.e =0.10; P <1.0×10^{-300}; r^2=0.20%).Mean corpuscular haemoglobin (MCH) showed association with white (β =0.27; s.e =5.94×10^{-3}; P <1.0×10^{-300}; r^2= 0.89%), and red wine (β =0.30; s.e =6.95×10^{-03}; P <1.0×10^{-300}; r^2=1.23%). Mean corpuscular volume (MCV) was associated with white wine (β =0.12; s.e =2.5×10^{-3}; P <1.0×10^{-300}; r^2=1.25%) and red wine (β =0.12; s.e =2.9×10^{-3}; P <1.0×10^{-300}; r^2=1.16%). Mean sphered cell volume (MCSV) was associated with white wine (β =0.09; s.e =2.05×10^{-03}; P <1.0×10^{-300}; r^2=1.04%).

Systolic blood pressure showed evidence of an association with beer/cider consumption (β=0.02; s.e = 6.0×10^{-4}; P <1.0×10^{-300}; r^2= 2.4%).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2022 doi:10.20944/preprints202205.0258.v1

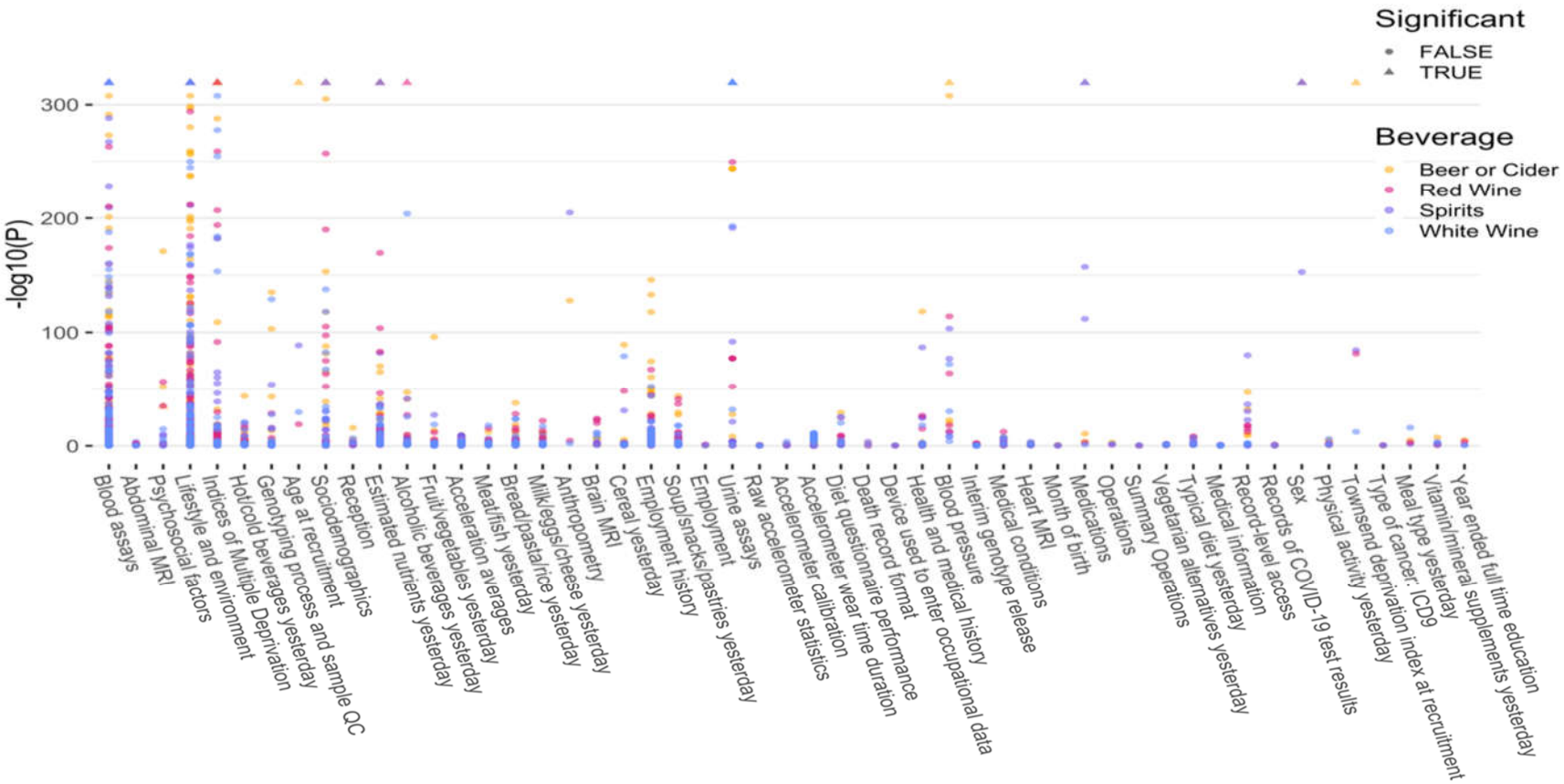

Figure 3. Overview of the agnostic association analyses. Manhattan plot illustrates negative log10 observed p values for general linear models created for association of UK biobank traits with consumption of various alcoholic beverages. UK biobank traits have been categorised into groups. Significant models are plotted as triangles and non-significant models are plotted as circles. Beer/cider results are plotted in yellow, red wine results are plotted in pink, spirit results are plotted in purple and white wine results are plotted in blue.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2022 doi:10.20944/preprints202205.0258.v1

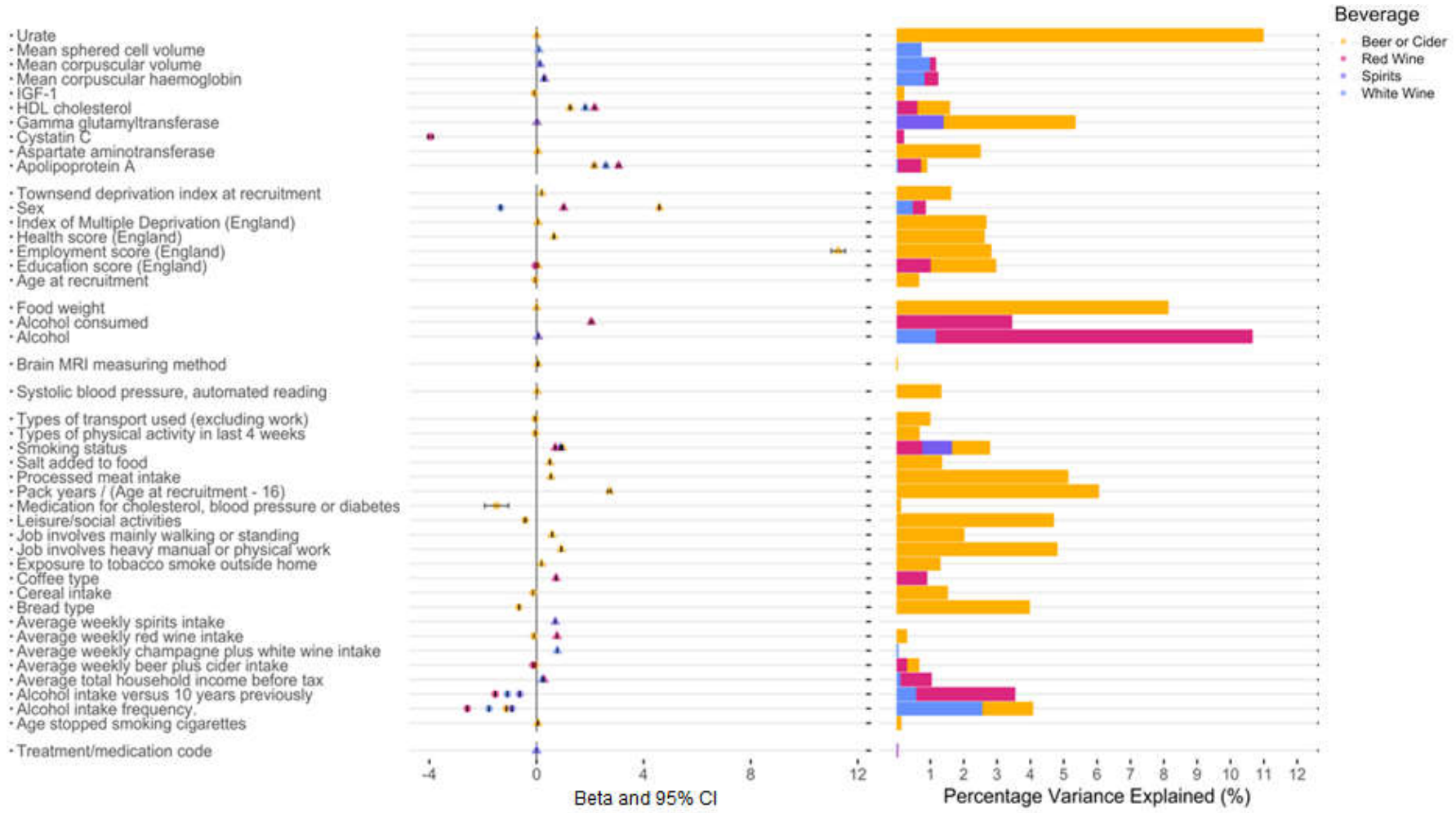

Figure 4. Overview of leading associations between consumption of alcoholic beverages and various traits within the UK biobank. A. The effect estimates and confidence intervals for the leading associations are depicted. **B**. Percentage variance explained for association of various alcoholic beverage consumption

 doi:10.20944/preprints202205.0258.v1

and UK biobank traits is depicted. Beer/cider results are plotted in yellow, red wine results are plotted in pink, spirit results are plotted in purple and white wine results are plotted in blue.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2022 doi:10.20944/preprints202205.0258.v1

In our MR analysis using IVW method, alcohol consumption showed a causal association with the liver enzyme GGT (β=9.72; CI=5.78-13.65; P_{emp}=5.61×10^{-5}; **Table 2**). We also observed significant causal association between alcohol consumption and multiple lifestyle choices including dietary factors such as individuals' preference to wholemeal or wholegrain (β=-0.054; CI=-0.09, -0.02; P_{emp}=5.69×10^{-3}) and white bread preference (β=0.054; CI,0.02-0.09; P_{emp}=1.99×10^{-3}). We observed significant causal association between alcohol consumption and MSCV (β=0.23; CI= 0.15-0.31; P_{emp}=3.70×10^{-6}**),** MCV (β=0.27; CI=0.18-0.36; P_{emp}=2.36×10^{-6}) and MCH (β=0.27; CI=0.18-0.36; P_{emp}=1.60×10^{-6}) yielded significant results.

Table 2. Overview of the results of Mendelian Randomization analysis.

	Single Instrument MR			Multiple Instrument MR		
	Beta	**95% CI**	**P value**	Beta	95% CI	P value
Gamma Glutamyl Transferase	9.7	5.8,13.6	1×10^{-4}	9.4	5.9, 12.9	0
Mean Sphered Cell Volume	0.2	0.15,0.31	0	0.19	0.11, 027	1 ×10^{-4}
Mean Corpuscular Haemoglobin	0.3	0.18, 0.36	0	0.3	0.19,0.35	0
Mean Corpuscular Volume	0.3	0.18, 0.36	0	0.3	0.19,0.36	0
Unplanned physical activity by method of transport	-0.04	-0.07, -0.01	0.01	-0.06	-0.1, -0.03	8 ×10^{-4}
Wholemeal/ wholegrain bread consumption	-0.05	-0.09, -0.02	0.006	-0.06	-0.1, -0.02	0.008
White bread consumption	0.05	0.02, 0.09	0.002	0.05	0.02, 0.09	0.008

The results from PheWAS analysis (**Table 3)** showed significant association between causal alcohol SNP (rs1229984) and alcohol-relating disorders (β =0.237; s.e=0.038; P =4.78×10^{-10}), alcohol dependency (β =0.264; s.e=0.047; P =2.52×10^{-8}), alcoholic liver damage (β =0.271; s.e=0.058; P =3.47×10^{-6}), and enthesopathy (β =-0.064; s.e=0.014; P =1.05×10^{-5}).

Table 3. Overview of the rs1229984 PheWAS results in the UK biobank cohort.

Description	Effect Estimate	Standard Error	Odds Ratio	P-Value
Alcohol-related disorders	0.24	0.04	1.3	4.87 × 10^{-10}
Alcoholism	0.26	0.05	1.3	2.52 × 10^{-08}
Alcoholic liver damage	0.27	0.06	1.3	3.47 × 10^{-06}
Enthesopathy	-0.06	0.01	0.9	1.05 × 10^{-05}

In our mediation analysis, GGT showed mediating the effect of alcohol consumption on alcohol dependence (β=0.016; se=0.0063; P=0.0147). In addition, physical inactivity mediated the effect of alcohol consumption on alcoholic liver cirrhosis (β=-8.96; se=1.38; P=2.91×10^{-3}) (**Figure 5**).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2022 doi:10.20944/preprints202205.0258.v1

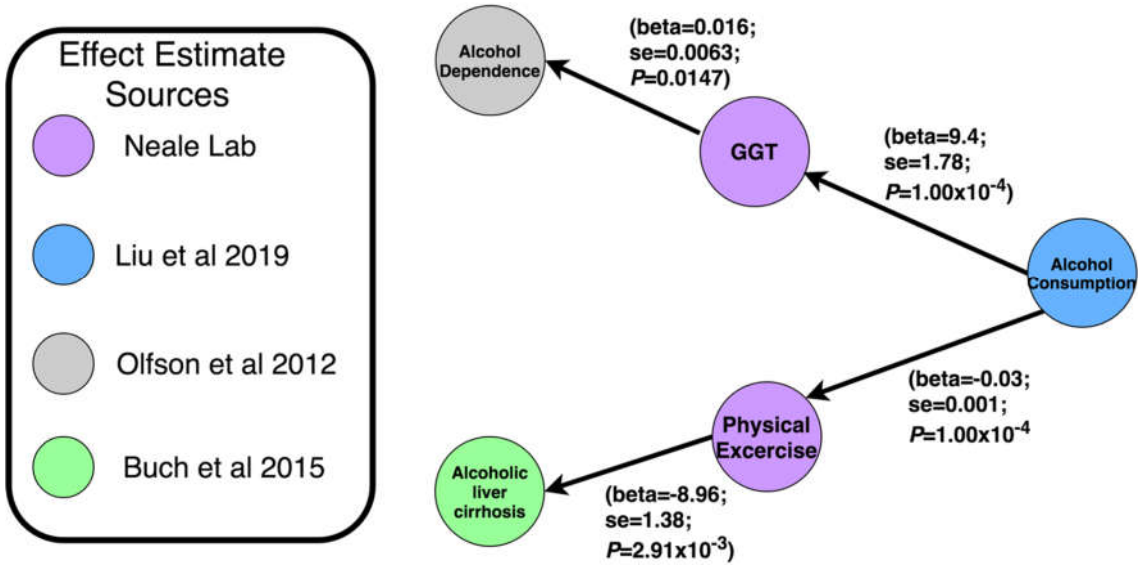

Figure 5. Summary of the mediation analysis. Network to summarise mediators of alcohol consumption and alcohol related diseases. Each node represents a trait that was included in mediation analysis. Results from MR analysis between each pair of traits is depicted on the edges. Data source to obtain summary statistics for each trait is depicted in the legend. Beta: Effect estimate from Mendelian Randomization; Se: standard error; P: P-value; GGT: Gamma Glutamyl transferase.

 doi:10.20944/preprints202205.0258.v1

Discussion

Here, we found evidence of 20 causal relationships between UK biobank phenotypes and alcohol consumption. We also identified (1) GGT as a possible mediator of alcohol consumption's effect on alcohol dependence and (2) physical activity as a possible mediator of alcohol liver cirrhosis. The identification of these risk factors between alcohol consumption and various diseases may help identification of individuals at higher risk of developing alcohol-related diseases.

We observed that the liver enzyme GGT showed a significant association with alcohol consumption. This analysis found that individuals who consume large amounts of spirit and beer or cider per week tend to have increased serum GGT. This association was further supported by the MR analyses which provided evidence of causal relationships between alcohol consumption and GGT. Furthermore, we identified GGT as a possible mediator of alcohol consumption and alcohol dependence. This finding is in line with previous epidemiological evidence and goes further to show the causality of this relationship. A four-year prospective study with 6846 male participants observed that alcohol consumption was associated with a raised blood levels of multiple liver enzymes, including GGT (Duk-Hee Lee, Myung-Hwa Ha, & David C. Christiani, 2001). Furthermore, one study found that moderate alcohol drinking (which was defined as less than 40g ethanol per day) raised GGT but did not significantly raise other liver enzymes such as AST (Päivikki I. Alatalo et al., 2008). It was previously shown that BMI has larger effect on liver enzymes than alcohol consumption alone (Duk-Hee Lee et al., 2001). Our analysis, which included excessive drinkers and was corrected for BMI, aligns with these epidemiological findings, and additionally demonstrate a causal relationship between beer/cider and spirit consumption with GGT.

We identified GGT and physical exercise as mediators of alcohol related diseases. A previous MR study found evidence of positive causal relationships between alcohol consumption, BMI and liver enzymes such as GGT (A. R. Carter et al., 2019). Our analysis highlights physical activity as a mediator of alcoholic liver cirrhosis. This could be possibly due to the impact that physical activity has on BMI and therefore non-alcoholic fatty liver disease as an important source for liver cirrhosis.

Our initial agnostic analysis showed associations between erythrocyte phenotypes and alcohol consumption. Our MR analyses showed a positive causal relationship between wine consumption and erythrocyte characteristics. Grape and wine products contain substantial amount of iron (S. Galani-Nikolakaki, N. Kallithrakas-Kontos, & A. A. Katsanos, 2002) that is a fundamental trace element in production of erythrocytes. In our study, alcohol consumption was linked to increased MSCV, MCV, and MCH. Severe alcoholism has been associated with anaemia and raised reticulocyte count and size (M. Myrhed, L. Berglund, & L. E. Böttiger, 1977) (H. S. Ballard, 1997). Toth et al. showed increased hematologic parameters in 39 healthy non-smoking volunteers after exposure to red wine (A. Toth et al., 2014). Non-alcoholic properties of red wine have been shown to act as antioxidant (I. Tedesco et al., 2000).

Our study highlighted that beer/cider consumption was linked to serum urate. Excess alcohol consumption is well documented to be linked to hyperuricemia (E. W. Campion, R. J. Glynn, & L. O. DeLabry, 1987) (Y. H. Jee, K. J. Jung, Y. B. Park, W. Spiller, & S. H. Jee, 2019). Specifically, beer and spirits have been shown to be linked to increases serum urate

 doi:10.20944/preprints202205.0258.v1

compared with other types of alcoholic drinks (H. K. Choi & G. Curhan, 2004; Angelo L. Gaffo et al., 2010).

Our analysis benefited from the large sample size and the rich phenotyping of the UK biobank cohort (Cathie Sudlow et al., 2015) and many MR studies we performed to investigate the causality of these associations. This increased our statistical power to detect traits associated with alcoholic beverage consumption (Xue Li et al., 2018). Furthermore, our beverage-specific agnostic approach reduces bias in our results, ensuring identification of novel traits associated with alcohol consumption (Clare Bycroft et al., 2018). Finally, mediation analysis gave us a better understanding of the relationships between alcohol associated traits and the alcohol related diseases (Eleanor Sanderson, 2021).

Our MR analysis was limited by the lack of beverage specific genetic instruments (e.g. instruments specific to beer/cider or spirit consumption). This would have allowed us to test for causal links specific to these beverages and not be limited to instruments associated with general alcohol consumption. One possible reason for this limitation could be that genetic components have been shown to account for a small amount of variance in alcohol consumption (Michelle Taylor, Andrew J. Simpkin, Philip C. Haycock, Frank Dudbridge, & Luisa Zuccolo, 2016). In our study, we had 34 self-reported traits. Self-reported data (*e.g.* physical exercise) are less generalizable compared to a measured phenotype (*e.g.* BMI) due to recall bias. Our conservative approach in performing multiple stage of analyses starting from agnostic association analysis, MR analysis, PheWAS and mediation analysis that made use of various data sources ensures improvement in accuracy of the results presented.

Conclusion

We took an agnostic approach to identify the causal genetic factors associated with alcoholic beverage consumption. Our analysis identified traits, a liver enzyme and diseases with direct or mediated causal links with increased alcohol consumption. Our findings also imply that GGT may mediate the effect of alcohol dependence.

Acknowledgments and Sources of Funding: R.P. holds a fellowship supported by Rutherford Fund from Medical Research Council (MR/R0265051/1 and MR/R0265051/2). F.O. is supported by Rutherford Fund from Medical Research Council MR/R0265051/2. X.J. is supported by Rutherford Fund from Medical Research Council MR/R0265051/2. SH.A. is supported by Rutherford Fund from Medical Research Council MR/R0265051/2.

References

Alatalo, Päivikki I., Koivisto, Heidi M., Hietala, Johanna P., Puukka, Katri S., Bloigu, Risto, & Niemelä, Onni J. (2008). Effect of moderate alcohol consumption on liver enzymes increases with increasing body mass index. *Am J Clin Nutr, 88*(4), 1097-1103. doi:10.1093/ajcn/88.4.1097

Alfaro-Almagro, Fidel, McCarthy, Paul, Afyouni, Soroosh, Andersson, Jesper L. R., Bastiani, Matteo, Miller, Karla L., . . . Smith, Stephen M. (2021). Confound modelling in UK Biobank brain imaging. *NeuroImage, 224*, 117002-117002. doi:10.1016/j.neuroimage.2020.117002

Ballard, H. S. (1997). The hematological complications of alcoholism. *Alcohol Health Res World, 21*(1), 42-52.

Bierut, L. J., Agrawal, A., Bucholz, K. K., Doheny, K. F., Laurie, C., Pugh, E., . . . Rice, J. P. (2010). A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci U S A, 107*(11), 5082-5087. doi:10.1073/pnas.0911109107

Buch, S., Stickel, F., Trépo, E., Way, M., Herrmann, A., Nischalke, H. D., . . . Hampe, J. (2015). A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. *Nat Genet, 47*(12), 1443-1448. doi:10.1038/ng.3417

Burgess, S., & Thompson, S. G. (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol, 40*(3), 755-764. doi:10.1093/ije/dyr036

Bycroft, Clare, Freeman, Colin, Petkova, Desislava, Band, Gavin, Elliott, Lloyd T., Sharp, Kevin, . . . Marchini, Jonathan. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature, 562*(7726), 203-209. doi:10.1038/s41586-018-0579-z

Campion, E. W., Glynn, R. J., & DeLabry, L. O. (1987). Asymptomatic hyperuricemia. Risks and consequences in the Normative Aging Study. *Am J Med, 82*(3), 421-426. doi:10.1016/0002-9343(87)90441-4

Carter, A. R., Borges, M. C., Benn, M., Tybjærg-Hansen, A., Davey Smith, G., Nordestgaard, B. G., & Lawlor, D. A. (2019). Combined Association of Body Mass Index and Alcohol Consumption With Biomarkers for Liver Injury and Incidence of Liver Disease: A Mendelian Randomization Study. *JAMA Netw Open, 2*(3), e190305. doi:10.1001/jamanetworkopen.2019.0305

Choi, H. K., & Curhan, G. (2004). Beer, liquor, and wine consumption and serum uric acid level: the Third National Health and Nutrition Examination Survey. *Arthritis Rheum, 51*(6), 1023-1029. doi:10.1002/art.20821

Clarke, Laura, Zheng-Bradley, Xiangqun, Smith, Richard, Kulesha, Eugene, Xiao, Chunlin, Toneva, Iliana, . . . Genomes Project, Consortium. (2012). The 1000 Genomes Project: data management and community access. *Nature methods, 9*(5), 459-462. doi:10.1038/nmeth.1974

Davey Smith, George, & Hemani, Gibran. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet, 23*(R1), R89-R98. doi:10.1093/hmg/ddu328

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and their Application*: Cambridge University Press.

Denny, J. C., Ritchie, M. D., Basford, M. A., Pulley, J. M., Bastarache, L., Brown-Gentry, K., . . . Crawford, D. C. (2010). PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics, 26*(9), 1205-1210. doi:10.1093/bioinformatics/btq126

Duell, Eric J., Sala, Núria, Travier, Noémie, Muñoz, Xavier, Boutron-Ruault, Marie Christine, Clavel-Chapelon, Françoise, . . . González, Carlos A. (2011). Genetic variation in alcohol dehydrogenase (ADH1A , ADH1B , ADH1C , ADH7) and aldehyde dehydrogenase (ALDH2), alcohol consumption and gastric cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. *Carcinogenesis, 33*(2), 361-367. doi:10.1093/carcin/bgr285

Edenberg, Howard J., & McClintick, Jeanette N. (2018). Alcohol Dehydrogenases, Aldehyde Dehydrogenases, and Alcohol Use Disorders: A Critical Review. *Alcoholism: Clinical and Experimental Research, 42*(12), 2281-2297. doi:https://doi.org/10.1111/acer.13904

Gaffo, Angelo L., Roseman, Jeffrey M., Jacobs, David R., Jr., Lewis, Cora E., Shikany, James M., Mikuls, Ted R., . . . Saag, Kenneth G. (2010). Serum urate and its relationship with alcoholic beverage intake in men and women: findings from the Coronary Artery Risk Development in Young Adults (CARDIA) cohort. *Annals of the rheumatic diseases, 69*(11), 1965-1970. doi:10.1136/ard.2010.129429

Galani-Nikolakaki, S., Kallithrakas-Kontos, N., & Katsanos, A. A. (2002). Trace element analysis of Cretan wines and wine products. *Science of The Total Environment, 285*(1), 155-163. doi:https://doi.org/10.1016/S0048-9697(01)00912-3

Hemani, Gibran, Zheng, Jie, Elsworth, Benjamin, Wade, Kaitlin H., Haberland, Valeriia, Baird, Denis, . . . Haycock, Philip C. (2018). The MR-Base platform supports systematic causal inference across the human phenome. *elife, 7*, e34408. doi:10.7554/eLife.34408

Jee, Y. H., Jung, K. J., Park, Y. B., Spiller, W., & Jee, S. H. (2019). Causal effect of alcohol consumption on hyperuricemia using a Mendelian randomization design. *Int J Rheum Dis, 22*(10), 1912-1919. doi:10.1111/1756-185x.13668

Jorgenson, E., Thai, K. K., Hoffmann, T. J., Sakoda, L. C., Kvale, M. N., Banda, Y., . . . Choquet, H. (2017). Genetic contributors to variation in alcohol consumption vary by race/ethnicity in a large multi-ethnic genome-wide association study. *Mol Psychiatry, 22*(9), 1359-1367. doi:10.1038/mp.2017.101

Kiiskinen, Tuomo, Mars, Nina J., Palviainen, Teemu, Koskela, Jukka, Rämö, Joel T., Ripatti, Pietari, . . . FinnGen, Gscan Consortium. (2020). Genomic prediction of alcohol-related morbidity and mortality. *Translational Psychiatry, 10*(1), 23. doi:10.1038/s41398-019-0676-2

Kyle, Simon D., Sexton, Claire E., Feige, Bernd, Luik, Annemarie I., Lane, Jacqueline, Saxena, Richa, . . . Spiegelhalder, Kai. (2017). Sleep and cognitive performance: cross-sectional associations in the UK Biobank. *Sleep Med, 38*, 85-91. doi:10.1016/j.sleep.2017.07.001

Lee, Duk-Hee, Ha, Myung-Hwa, & Christiani, David C. (2001). Body weight, alcohol consumption and liver enzyme activity—a 4-year follow-up study. *Int J Epidemiol, 30*(4), 766-770. doi:10.1093/ije/30.4.766

Li, Xue, Meng, Xiangrui, Spiliopoulou, Athina, Timofeeva, Maria, Wei, Wei-Qi, Gifford, Aliya, . . . Theodoratou, Evropi. (2018). MR-PheWAS: exploring the causal effect of SUA level on multiple disease outcomes by using genetic instruments in UK Biobank. *Annals of the rheumatic diseases, 77*(7), 1039-1047. doi:10.1136/annrheumdis-2017-212534

Liu, Mengzhen, Jiang, Yu, Wedow, Robbee, Li, Yue, Brazel, David M., Chen, Fang, . . . Vrieze, Scott. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature genetics, 51*(2), 237-244. doi:10.1038/s41588-018-0307-5

Myrhed, M., Berglund, L., & Böttiger, L. E. (1977). Alcohol consumption and hematology. *Acta Med Scand, 202*(1-2), 11-15. doi:10.1111/j.0954-6820.1977.tb16774.x

Peakman, T. C., & Elliott, P. (2008). The UK Biobank sample handling and storage validation studies. *Int J Epidemiol, 37 Suppl 1*, i2-6. doi:10.1093/ije/dyn019

Purcell, Shaun, Neale, Benjamin, Todd-Brown, Kathe, Thomas, Lori, Ferreira, Manuel A. R., Bender, David, . . . Sham, Pak C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet, 81*(3), 559-575. doi:10.1086/519795

Sanderson, Eleanor. (2021). Multivariable Mendelian Randomization and Mediation. *Cold Spring Harbor perspectives in medicine, 11*(2), a038984. doi:10.1101/cshperspect.a038984

Sudlow, Cathie, Gallacher, John, Allen, Naomi, Beral, Valerie, Burton, Paul, Danesh, John, . . . Collins, Rory. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med, 12*(3), e1001779-e1001779. doi:10.1371/journal.pmed.1001779

Taylor, Michelle, Simpkin, Andrew J., Haycock, Philip C., Dudbridge, Frank, & Zuccolo, Luisa. (2016). Exploration of a Polygenic Risk Score for Alcohol Consumption: A Longitudinal Analysis from the ALSPAC Cohort. *PLoS One, 11*(11), e0167360-e0167360. doi:10.1371/journal.pone.0167360

Tedesco, I., Russo, M., Russo, P., Iacomino, G., Russo, G. L., Carraturo, A., . . . Palumbo, R. (2000). Antioxidant effect of red wine polyphenols on red blood cells. *J Nutr Biochem, 11*(2), 114-119. doi:10.1016/s0955-2863(99)00080-7

Teslovich, Tanya M., Musunuru, Kiran, Smith, Albert V., Edmondson, Andrew C., Stylianou, Ioannis M., Koseki, Masahiro, . . . Kathiresan, Sekar. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature, 466*(7307), 707-713. doi:10.1038/nature09270

Toth, A., Sandor, B., Papp, J., Rabai, M., Botor, D., Horvath, Z., . . . Czopf, L. (2014). Moderate red wine consumption improves hemorheological parameters in healthy volunteers. *Clin Hemorheol Microcirc, 56*(1), 13-23. doi:10.3233/ch-2012-1640

Verbanck, Marie, Chen, Chia-Yen, Neale, Benjamin, & Do, Ron. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature genetics, 50*(5), 693-698. doi:10.1038/s41588-018-0099-7

Verma, Anurag, & Ritchie, Marylyn D. (2017). Current Scope and Challenges in Phenome-Wide Association Studies. *Current epidemiology reports, 4*(4), 321-329. doi:10.1007/s40471-017-0127-7

World Health, Organization. (2018). *Global status report on alcohol and health 2018: executive summary*. Retrieved from Geneva: https://apps.who.int/iris/handle/10665/312318

Zheng, Jie, Baird, Denis, Borges, Maria-Carolina, Bowden, Jack, Hemani, Gibran, Haycock, Philip, . . . Smith, George Davey. (2017). Recent Developments in Mendelian Randomization Studies. *Current epidemiology reports, 4*(4), 330-345. doi:10.1007/s40471-017-0128-6