Smoking, Alcohol Consumption, and 24 Gastrointestinal Diseases: Mendelian Randomization Analysis

Running head: Smoking, alcohol intake, and 24 gastrointestinal diseases

Shuai Yuan[1,2#], Jie Chen[1,3#], Xixian Ruan[3#], Yuhao Sun[1], Ke Zhang[4,5], Xiaoyan Wang[3*], Xue Li[1,6*], Dipender Gill[7], Stephen Burgess[8,9], Edward L Giovannucci[10,11], Susanna C. Larsson[2,12]

Affiliations

[1] School of Public Health and The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China.

[2] Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

[3] Department of Gastroenterology, The Third Xiangya Hospital, Central South University, Changsha, China.

[4] Key Laboratory of Growth Regulation and Translational Research of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, China.

[5] Westlake Intelligent Biomarker Discovery Lab, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, China.

[6] Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh, UK.

[7] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.

[8] MRC Biostatistics Unit, University of Cambridge, Cambridge, UK.

[9] Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

[10] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[11] Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[12] Unit of Medical Epidemiology, Department of Surgical Sciences, Uppsala University, Uppsala, Sweden.

[#] Shuai Yuan, Jie Chen, and Xixian Ruan share first authorship

[*] Xiaoyan Wang and Xue Li share corresponding authorship

Correspondence to

Xue Li, School of Public Health and the Second Affiliated Hospital, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China. E-mail: xue.li@ed.ac.uk

Xiaoyan Wang, Department of Gastroenterology, The Third Xiangya Hospital, Central South University, 138 Tongzipo Road, Changsha, Hunan 410013, P.R. China. E-mail: wangxiaoyan@csu.edu.cn

Word count: abstract, 252; text, 3801

Number of tables and figures: 2 tables and 2 figures

Abstract

Background: Whether the positive associations of smoking and alcohol consumption with gastrointestinal diseases are causal is uncertain. We conducted this Mendelian randomization (MR) to comprehensively examine associations of smoking and alcohol consumption with common gastrointestinal diseases.

Methods: Genetic variants associated with smoking initiation and alcohol consumption at the genome-wide significance level were selected as instrumental variables. Genetic associations with 24 gastrointestinal diseases were obtained from the UK Biobank, FinnGen study, and other large consortia. Univariable and multivariable MR analyses were conducted to estimate the overall and independent MR associations after mutual adjustment for genetic liability to smoking and alcohol consumption.

Results: Genetic predisposition to smoking initiation was associated with increased risk of 20 of 24 gastrointestinal diseases, including 7 upper gastrointestinal diseases (gastroesophageal reflux, esophageal cancer, gastric ulcer, duodenal ulcer, acute gastritis, chronic gastritis and gastric cancer), 4 lower gastrointestinal diseases (irritable bowel syndrome, diverticular disease, Crohn's disease and ulcerative colitis), 8 hepatobiliary and pancreatic diseases (non-alcoholic fatty liver disease, alcoholic liver disease, cirrhosis, liver cancer, cholecystitis, cholelithiasis, acute and chronic pancreatitis), and acute appendicitis. Fifteen out of 21 associations persisted after adjusting for genetically-predicted alcohol consumption. Genetically-predicted higher alcohol consumption was associated with increased risk of duodenal cancer, alcoholic liver disease, cirrhosis, and chronic pancreatitis; however, the association for duodenal ulcer did not remain after adjustment for genetic predisposition to smoking initiation.

Conclusion: This study provides MR evidence supporting causal associations of smoking with a broad range of gastrointestinal diseases, whereas alcohol consumption was associated with only a few gastrointestinal diseases.

Funding: The Natural Science Fund for Distinguished Young Scholars of Zhejiang Province; National Natural Science Foundation of China; Key Project of Research and Development Plan of Hunan Province; the Swedish Heart Lung Foundation; the Swedish Research Council; the Swedish Cancer Society.

Keywords: alcohol consumption; gastrointestinal diseases; Mendelian randomization; smoking

Introduction

Tobacco smoking and alcohol consumption are leading causes of the global burden of disease and are major contributors to premature mortality (1, 2). Gastrointestinal diseases account for considerable health care use and expenditures, and a holistic approach to lifestyle interventions may result in more health gains and less economic burdens (3). Population-based studies have identified tobacco smoking as a risk factor for several gastrointestinal diseases, including gastroesophageal reflux disease (4), esophageal cancer (5), Crohn's disease (6), liver cancer (7), and pancreatitis (8). Evidence on the association between tobacco smoking and risk of other gastrointestinal diseases is limited and inconsistent. Like smoking, heavy alcohol consumption has been associated with increased risk of several gastrointestinal outcomes, including gastritis (9), gastric cancer (10), colorectal cancer (11), cirrhosis (12), liver cancer (7), and pancreatitis (8). However, whether these associations are all causal remain unestablished since most of the evidence was obtained from observational studies where the results may be biased by reverse causality and confounding. Of note, even though reverse causality may not be an issue in the studies for any of studied gastroenterological outcomes, it might exist for certain gastroenterological diseases causing pain, which smoker patients may try to increase smoking dose to mitigate via an intake of higher levels of nicotine. In addition, as smoking and alcohol consumption are phenotypically and genetically correlated (13, 14), the independent impacts of smoking and alcohol consumption on gastrointestinal

diseases are unclear. Establishing the causal association of tobacco smoking and alcohol consumption with gastrointestinal diseases is crucial, as this provides further evidence for subsequent recommending public policies and clinical interventions.

Mendelian randomization (MR) is an epidemiological approach that utilizes genetic variants as an instrument to strengthen the causal inference in an exposure-outcome association (15). MR is by nature not prone to confounding since genetic variants are randomly assorted at conception and thus unrelated to environmental and self-adopted factors that usually act as confounders. Additionally, this method can minimize reverse causality since fixed alleles are unaffected by the onset and progression of disease. Previous MR studies have examined the associations of smoking and alcohol consumption with several gastrointestinal diseases (16-21). Nevertheless, whether smoking and alcohol consumption exert influence on a wide range of gastrointestinal outcomes have not been investigated in a comprehensive way. A thorough investigation on the gastrointestinal consequences of smoking and alcohol drinking is of great importance to develop non-pharmacological interventions on gastrointestinal diseases. Here, we conducted an MR investigation of the associations of smoking and alcohol consumption with the risk of common gastrointestinal diseases to fill in above knowledge gaps.

Materials and Methods

Fig 1 shows the study design overview. The study was based on publicly available genome-wide association studies (GWAS), and the detailed information on used studies was presented in **S1 Table**. The genetic associations were estimated using data from the UK Biobank study (22), the FinnGen study (23), and several large consortia. The summary effect estimates were combined using meta-analysis for each gastrointestinal disease from different data resources. Included studies had been approved by corresponding institutional review boards and ethical committees, and consent forms had been signed by all participants.

Instrumental variable selection

A total of 378 and 99 single nucleotide polymorphisms (SNPs) associated with smoking initiation (a binary phenotype indicating whether an individual had ever being a regular smoker, 1,232,091 individuals of European descent) and alcohol consumption (log-transformed drinks per week, 941,280 individuals of European descent) at the genome-wide significance threshold ($P<5\times10^{-8}$) were identified by the GWAS and Sequencing Consortium of Alcohol and Nicotine use (GSCAN) study (14). These SNPs explained approximately 2.3% and 0.3% of the variation in smoking initiation and alcohol consumption, respectively (14). SNPs in linkage disequilibrium (defined as $r^2>0.01$ or clump distance <10,000 kb) and with the weaker associations with the exposure were removed, leaving 314 independent SNPs as instrumental variables for smoking initiation and 84 for alcohol consumption. Smoking initiation and

alcohol consumption shared two index genetic variants, which were rs1713676 and rs11692435. Considering partial sample overlap (around 30%) between the GSCAN study with full data and the UK Biobank study (14), we performed sensitivity analyses for smoking initiation and alcohol consumption using summary statistics data from the analysis excluding the UK Biobank and 23andMe. For a supplementary analysis of smoking behavior, we used 126 SNPs associated with a lifetime smoking index that considered smoking duration, heaviness, and cessation (24). The set of genetic instruments captured around 0.36% of the variance in lifetime smoking (24). We also conducted a sensitivity analysis using rs1229984 in *ADH1B* gene that encodes alcohol dehydrogenase 1B enzyme as the genetic instrument for alcohol consumption to minimize pleiotropy. Detailed information on used SNPs is presented in **S2 Table**.

Gastrointestinal disease data sources

Genetic associations with 24 gastrointestinal diseases were obtained from the UK Biobank study (22), the FinnGen study (23), and two large consortia, including the International Inflammatory Bowel Disease Genetics Consortium (IIBDGC) (25) and Genetic Epidemiology Research on Aging (GERA) (26). Included outcomes were classified into four major categories according to the disease onset site: 1) upper gastrointestinal diseases (gastroesophageal reflux disease, esophageal cancer, gastric ulcer, acute gastritis, chronic gastritis, and gastric cancer); 2) lower gastrointestinal diseases (irritable bowel disease, celiac disease, diverticular disease,

Crohn's disease, ulcerative colitis, and colorectal cancer); 3) hepatobiliary and pancreatic disease (non-alcoholic fatty liver disease, alcoholic liver disease, cirrhosis, liver cancer, cholangitis, cholecystitis, cholelithiasis, acute pancreatitis, chronic pancreatitis, and pancreatic cancer); 4) other (acute appendicitis).

The UK Biobank study is a large multicenter cohort study of 500,000 participants recruited in the United Kingdom between 2006 and 2010 (22). We used the summary statistics of European ancestry from GWAS conducted by Lee lab where the gastrointestinal outcomes were defined by codes of the International Classification of Diseases 9th Revision (ICD-9) and ICD-10 (27). Genetic associations were estimated by logistic regression with adjustment for sex, birth year, and the first four genetic principal components. For the FinnGen study, we used summary-level data on the genetic associations with gastrointestinal diseases from the last publicly available R7 data release (23). The FinnGen study is a nationwide genetic study where genetic and electronic health record data were collected. The gastrointestinal diseases were ascertained by the codes of the ICD-8, ICD-9, and ICD-10. Genome-wide association analyses were adjusted for sex, age, genetic components, and genotyping batch. Summary-level genetic data on Crohn's disease (5,956 cases and 14,927 controls) and ulcerative colitis (6,968 cases and 20,464 controls) were additionally obtained from the IIBDGC (25) and data on irritable bowel syndrome (3,117 cases and 53,520

controls) were obtained from the GERA (26). Detailed diagnostic codes are listed in **S3 Table.**

Statistical analysis

Data were harmonized to omit ambiguous SNPs with non-concordant alleles and palindromic SNPs with ambiguous minor allele frequency (>0.42 and <0.58) were removed from the analysis. The primary MR analyses were performed by the multiplicative random-effects inverse-variance weighted (IVW) method, which provides the most precise estimates though assuming that all SNPs are valid instruments. The analysis of rs1229984 for alcohol consumption was conducted by the Wald method. Estimates for each association from different sources were combined using fixed-effects meta-analysis and the heterogeneity of the associations from different data sources were evaluated by the I^2 statistic. Heterogeneity among SNPs' estimates in each association was assessed by Cochran's Q value. Multivariable MR analyses were conducted to mutually adjust for smoking initiation and alcohol consumption. To detect potential unbalanced pleiotropy (horizontal pleiotropy) and examine the consistency of the associations, three sensitivity analyses including the weighted median (28), MR-Egger (29), and Mendelian randomization pleiotropy residual sum and outlier (MR-PRESSO) (30) analyses were performed. The weighted median method can provide consistent estimates when more than 50% of the weight comes from valid instrument variants (28). The

MR-Egger intercept test can detect unmeasured pleiotropy and MR-Egger regression can generate estimates after accounting for horizontal pleiotropy albeit with less precision (29). The MR-PRESSO method can identify SNP outliers and provide results identical to that from IVW after removal of outliers (30). The *F*-statistic was estimated to quantify instrument strength and an *F*-statistic >10 suggested a sufficiently strong instrument. Power analysis was performed using an online tool (31). The Benjamini-Hochberg correction that controls the false discovery rate was applied to correct for multiple testing. The association with a nominal *P*-value <0.05 but Benjamini-Hochberg adjusted *P*-value >0.05 was regarded suggestive and the association with a Benjamini–Hochberg adjusted *P*-value <0.05 were deemed significant. All analyses were two-sided and performed using the TwoSampleMR (32), MendelianRandomization (28), and MRPRESSO R packages (30) in R software 4.1.2.

Results

The *F*-statistic for each genetic variant was above 10, suggesting a good strength of used genetic instruments (**S2 Table**). Most associations were well powered (**S4 Table**). For smoking initiation, there was 80% power to detect the smallest odds ratio (OR) ranging from 1.08 to 1.40 for included outcomes. Although power was lower for alcohol consumption, it was adequate to detect a moderate effect size for most common gastrointestinal diseases.

224

225 **Smoking and gastrointestinal diseases**

226 Genetic predisposition to smoking initiation was associated with 20 of the 24 studied
227 gastrointestinal diseases and all these associations remained after multiple
228 comparison correction (**Table 1 and S5 Table**). In detail, genetic liability to smoking
229 initiation was positively associated with 7 upper gastrointestinal diseases:
230 gastroesophageal reflux (OR, 1.28; 95% confidence interval [CI], 1.20-1.37; P =
231 4.09×10^{-14}), esophageal cancer (OR, 1.67; 95% CI, 1.24-2.25; $P = 6.84\times10^{-4}$), gastric
232 ulcer (OR, 1.54; 95% CI, 1.37-1.72; $P = 3.83\times10^{-14}$), duodenal ulcer (OR, 1.53; 95%CI,
233 1.34-1.75; $P = 8.47\times10^{-10}$), acute gastritis (OR, 1.29; 95%CI, 1.09-1.53; $P = 0.003$),
234 chronic gastritis (OR, 1.33; 95%CI, 1.18-1.49; $P = 1.55\times10^{-6}$) and gastric cancer (OR,
235 1.42; 95%CI, 1.13-1.79; $P = 0.002$); genetic liability to smoking initiation was positively
236 associated with 4 lower gastrointestinal diseases: irritable bowel syndrome (OR, 1.22;
237 95%CI, 1.12-1.32; $P = 3.50\times10^{-6}$), diverticular disease (OR, 1.25; 95%CI, 1.18-1.33;
238 $P = 5.23\times10^{-14}$), Crohn's disease (OR, 1.25; 95%CI, 1.11-1.40; $P = 3.03\times10^{-4}$) and
239 ulcerative colitis (OR, 1.15; 95%CI, 1.04-1.26; $P = 0.004$); genetic liability to smoking
240 initiation was positively associated with 8 hepatobiliary and pancreatic diseases:
241 non-alcoholic fatty liver disease (OR, 1.49; 95%CI, 1.26-1.76; $P = 3.82\times10^{-6}$),
242 alcoholic liver disease (OR, 1.99; 95%CI, 1.65-2.41; $P = 1.49\times10^{-12}$), cirrhosis (OR,
243 1.68; 95%CI, 1.40-2.02; $P = 3.39\times10^{-8}$), liver cancer (OR, 1.57; 95%CI, 1.13-2.17; P =
244 0.007), cholecystitis (OR, 1.47; 95%CI, 1.29-1.68; $P = 4.71\times10^{-9}$), cholelithiasis (OR,

1.20; 95%CI, 1.13-2.27; $P = 5.75\times10^{-9}$), acute pancreatitis (OR, 1.39; 95%CI, 1.23-1.56; $P = 6.71\times10^{-8}$) and chronic pancreatitis (OR, 1.38; 95% CI, 1.17-1.64; $P = 1.79\times10^{-4}$); genetic liability to smoking initiation was positively associated with acute appendicitis (OR, 1.15; 95% CI, 1.08-1.23; $P = 1.27\times10^{-5}$). Results were consistent in sensitivity analyses. An indication of horizontal pleiotropy was observed in the analysis of esophageal cancer in the FinnGen study (*P* for MR-Egger intercept <0.05, **S6 Table**). Although MR-PRESSO detected 1 to 3 outliers, the associations persisted and remained significant after removal of these out-lying SNPs (**S6 Table**). When using the genetic variants for smoking initiation based on data without the UK Biobank and 23andMe studies, the associations attenuated slightly albeit remained significant after multiple comparisons (Fig S1 and Table S7). All associations were replicated in the supplementary analysis of the lifetime smoking index (**S7 Table**). After correcting for multiple testing, genetically predicted lifetime smoking index was significantly associated with 17 of 24 gastrointestinal diseases, where the patterns of associations were generally similar to the analysis for smoking initiation (**Fig S2 and S7 Table)**. In distinction to the analysis of smoking initiation, genetically predicted lifetime smoking index was not significantly associated with acute gastritis, gastric cancer, Crohn's disease, and ulcerative colitis, whereas genetically predicted lifetime smoking index was significantly associated with pancreatic cancer (OR, 2.09; 95% CI, 1.30-3.36).

In multivariable MR analysis adjusted for genetically predicted alcohol consumption, the associations between genetically predicted smoking initiation and gastrointestinal

diseases were consistent with that from univariable MR analysis (**Table 1 and S8 Table**). However, the associations became stronger with wider CIs, in particular the associations for gastrointestinal reflux, esophageal cancer, gastric ulcer, irritable bowel syndrome, diverticular disease, non-alcoholic fatty liver disease, alcoholic liver disease, and cholecystitis (**Table 1**). In addition, the association for pancreatic cancer became suggestive significant from null.

Alcohol consumption and gastrointestinal diseases

Genetically-predicted alcohol consumption was nominally positively associated with esophageal cancer (OR, 2.86; 95% CI, 1.18-6.91; P = 0.020), duodenal ulcer (OR, 1.92; 95% CI, 1.23-3.00; P = 0.004), alcoholic liver disease (OR, 14.35; 95% CI, 7.69-26.81; $P = 6.32\times10^{-17}$), cirrhosis (OR, 2.96; 95% CI, 1.50-5.85; P = 0.002) and chronic pancreatitis (OR, 2.96; 95% CI, 1.80-4.89; $P = 2.13\times10^{-5}$), and nominally inversely associated with irritable bowel disease (OR, 0.73; 95% CI 0.57-0.93; P = 0.012) (**Table 2**). After Benjamini–Hochberg correction, the associations for duodenal cancer, alcoholic liver disease, cirrhosis, and chronic pancreatitis remained (**S5 Table**). Results were consistent in sensitivity analyses, and no horizontal pleiotropy was detected (**S9 Table**). One outlier was detected in the analysis of duodenal ulcer in the FinnGen study, and the association slightly changed after removal of this outlier (**S9 Table**). Results were consistent in the sensitivity analysis where the genetic associations with alcohol consumption were obtained from the genome-wide

association analysis excluding the UK Biobank and 23andMe studies (**Fig S3 and S7 Table**). The associations were directionally consistent albeit with wider CIs in the analysis where alcohol consumption was instrumented by rs1229984 (**S10 Table**). The associations for alcoholic liver disease, cirrhosis, and chronic pancreatitis persisted after adjustment for genetic liability to smoking initiation and multiple testing correction (**Table 2 and S8 Table**).

Discussion

We conducted a comprehensive MR investigation to examine the causal role of smoking and alcohol consumption in 24 gastrointestinal diseases and the result summary of this comprehensive analysis is shown in **Fig 2** and **S11 Table**. We found robust associations between genetic predisposition to smoking and increased risk of 21 gastrointestinal outcomes independent of alcohol consumption, showing an extensive impact on gastrointestinal health. In contrast, genetically predicted alcohol consumption was robustly and predominantly associated with increased risk of liver and pancreatic diseases, including alcoholic liver disease, cirrhosis, and chronic pancreatitis after adjustment for smoking.

Corroborating and extending the previous observational studies, our MR investigation strengthened the evidence that smoking has a detrimental effect on gastrointestinal health and increases the risk of a broad range of gastrointestinal diseases, including

gastroesophageal reflux disease (4), esophageal cancer (33), gastric and duodenal ulcer (34), gastritis (35), gastric cancer (36), irritable bowel syndrome (37), diverticular disease (38), Crohn's disease (6), cirrhosis (39), liver cancer (7), cholelithiasis (40), acute and chronic pancreatitis (41), and acute appendicitis (42). In line with previous MR studies, the current MR study also found that smoking was associated with increased risk of gastroesophageal reflux disease (16), esophageal cancer (17), gastric cancer (17), diverticular disease (18) non-alcoholic fatty liver disease (19), cholelithiasis (20), and acute and chronic pancreatitis (21). As for ulcerative colitis, traditional observational studies revealed a decreased risk among current smokers (6, 43); however, a recent MR analysis including 12,366 ulcerative colitis cases did not verify this inverse association in the analysis where smoking initiation was instrumented by 363 SNPs (44). Based on data from 3 independent populations, our study provided genetic evidence that smoking was a causal risk factor for ulcerative colitis in the analysis including 16,770 cases. Observational studies found that smoking was associated with an increased risk of colorectal cancer in a dose-dependent manner (45), whereas the positive association was not observed in an MR analysis (17). The current study was in line with the above MR study and found no strong association between smoking initiation and colorectal cancer risk. Nevertheless, a previous MR analysis with a 52,775 colorectal cancer cases found that genetic prediction to lifetime smoking index was positively associated with risks of colorectal cancer (46), which might imply that our null finding might be caused by

insufficient power due to a relatively small sample size. Smoking has been identified as a well-established risk factor for pancreatic cancer (47). Interestingly, despite a null finding on the association of genetic liability to smoking initiation and pancreatic cancer in univariable MR analysis, the association became stronger and suggestively significant after adjusting for genetically predicted alcohol consumption. This might be explained by an inverse association between moderate alcohol consumption and pancreatic cancer. In addition, an adverse effect of smoking on pancreatic cancer was observed when using a smoking index as genetic instrument for lifetime smoking exposure. Our findings also provide novel evidence on the associations of smoking with the higher risk of cholecystitis and alcoholic liver disease independently of alcohol consumption, which need to be verified.

The pathogenic role of alcohol in alcoholic liver disease is well-established and was confirmed also in our MR analysis. Our MR evidence along with previous observational studies also supported alcohol consumption as a risk factor for esophageal cancer (48), cirrhosis (49), and chronic pancreatitis (50). Noteworthy, the association between alcohol consumption and esophageal cancer became positively nonsignificant in multivariable MR, which possibly explained by the synergistic effect of alcohol and smoking. However, the association between alcohol consumption and duodenal ulcer has been scarcely studied. A meta-analysis including a small number of studies with relatively small sample sizes indicated that alcohol consumption was

not associated with duodenal ulcer (51). This null finding is likely due to insufficient power. Alcohol drinking has been associated with increased risk of gastric, colorectal, and liver cancer as well as acute pancreatitis (52). These associations were not supported by our MR study. A possible explanation for this inconsistent findings is that heavy alcohol drinking is commonly associated with an unhealthy lifestyle and meager nutrition (53), which might exert confounding effects that could not be ruled out in previous observational studies. Another possible reason is that the U-shaped association could not be detected in MR analysis. For example, light drinking may be associated with decreased risk of these diseases (11). In addition, it is also possible that the null associations observed in present study might be a consequence of inadequate power given SNPs used to mimic alcohol consumption explained a small phenotypic variance. In agreement with previous studies, our MR investigation demonstrated no associations of alcohol consumption with the development of gastroesophageal reflux, Crohn's disease, or ulcerative colitis (4, 6, 44).

Many mechanisms have been proposed to support the observed positive associations between smoking and gastrointestinal diseases. Tobacco smoking has been shown to augment the production of numerous pro-inflammatory cytokines and decrease the levels of anti-inflammatory cytokines (54), which might mediate a variety of inflammation-associated gastrointestinal diseases. In addition, smoking may also generate impacts on the immune system, including inhibition of the function of

circulatory dendritic cells (55) and alteration signaling of Toll-like receptors (56), which might contribute to the autoimmune disease and occurrence of neoplasm. The underlying mechanisms behind the associations of alcohol consumption with gastrointestinal diseases have not been fully understood. In addition to direct mucosal damage, the metabolites of ethanol are accountable for a part of the inflammation of alcohol drinking on the liver (57) and the gastrointestinal tract (58).

This study investigated the impacts of smoking and alcohol consumption on a wide range of gastrointestinal disease. Based on our findings, promoting public awareness of the adverse impacts of tobacco smoking and alcohol consumption on gastrointestinal diseases is of particular importance and should be used as prevention strategies to lower gastrointestinal disease burden because these two factors are modifiable behavioral factors as possible targets of the pharmacal (59) and behavioral interventions. In addition, our results may help facilitate the guidelines of gastrointestinal disease prevention and the management of certain patients who have a subsequent high risk of gastrointestinal disease., like those with obesity and diabetes (60, 61).

The major strength of the present study is MR design, which minimized bias from confounding and reverse causality and thus improved the causal inference in the associations of smoking and alcohol consumption with gastrointestinal diseases. We

also used several independent outcome sources and combined the estimates, which increased statistical power as well as strengthened our findings by the observed consistency of results. Another strength is that we confined our analysis within the individuals of European ancestry, which minimized the population stratification bias.

This study also has several limitations. A major limitation of MR design is horizontal pleiotropy, which means that the used SNPs exerts effects on the outcomes not via the exposure but via alternative pathways. However, in this study, the bias caused by pleiotropic effects should be minimal since we observed no indications of horizontal pleiotropy in MR-Egger analysis, consistent results from a series of sensitivity analyses, and robust associations from multivariable MR analysis with mutual adjustment. Another limitation is the relatively small phenotypic variance of alcohol consumption (approximately 0.2%), which resulted in inadequate power to detect weak associations for certain uncommon gastrointestinal diseases. There are several limitations of using summary-level data. First, we could not evaluate the nonlinear associations between alcohol consumption and gastrointestinal diseases without individual-level data. We could not differentiate the associations of smoking and alcohol consumption on the pathological subtypes of certain gastroenterological diseases, like esophageal cancer, based on summary-level data. For example, heavy alcohol consumption was associated with a high risk of squamous esophageal cancer (62), but the associations were inconsistent for adenocarcinoma esophageal cancer

(63), which needs further investigation. Stratification analysis on sex was unlikely to be performed. In addition, we could not interpret and rescale the associations in a comparable scale to traditional observational studies because the unit of the exposure phenotypes were fixed in the corresponding genome-wide association analyses. An additional limitation is that our analysis was confined to the European populations, and thus whether the observed associations can be generalized to other populations remains unknown. For alcohol consumption, it has been reported that there were substantial behavioral and genetic differences across ethnic groups. For example, East Asian individuals drink much less alcohol compared to other races, which appears to be related to *ALDH2* gene (64). A further potential limitation is that the UK Biobank study were included in both the exposure and outcome datasets, which might cause MR estimates towards the observational associations. However, the used instrumental variants were proven to be strongly associated with the exposure (*F*-statistic > 10) (65), and the associations were replicated in the FinnGen study. Moreover, the associations remained stable in the sensitivity analyses using the genetic associations with exposures from the data excluding the UK Biobank and 23andMe studies. All of these indicated that the bias due to sample overlap was limited.

In conclusion, this MR study suggested that smoking is a risk factor for a broad range of gastrointestinal diseases independent of alcohol consumption. Alcohol

consumption on the other hand seemed to be an independent risk factor for only a few gastrointestinal diseases, including alcoholic liver disease, cirrhosis, and chronic pancreatitis, but we cannot rule out weak associations with other diseases. These findings provide genetic evidence on supporting reducing tobacco smoking and possibly excessive alcohol consumption in particular to prevent gastrointestinal diseases.

Declarations

Availability of data and materials

Data analyzed in the current study are publicly available GWAS summary-level data. The specific information and link could be found in Table S1. The code and curated data for the current analysis are available at https://github.com/XixianRuan/smoking_gi.

Funding

XL: the Natural Science Fund for Distinguished Young Scholars of Zhejiang Province (LR22H260001). XYW: National Natural Science Foundation of China (81970494) and Key Project of Research and Development Plan of Hunan Province(2019SK2041); SCL: the Swedish Heart Lung Foundation (Hjärt-Lungfonden, 20210351), the Swedish Research Council (Vetenskapsrådet, 2019-00977), and the Swedish Cancer Society (Cancerfonden). Funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Acknowledgement

We want to thank the Lee Lab, the FinnGen study, the International Inflammatory Bowel Disease Genetics Consortium (IIBDGC), and the Genetic Epidemiology Research on Aging (GERA) for sharing data.

Conflict of interest

All authors declare no competing interest.

References

1. Collaborators GBDA. Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2018;392(10152):1015-35.

2. Collaborators GBDRF. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1223-49.

3. Peery AF, Crockett SD, Murphy CC, Jensen ET, Kim HP, Egberg MD, et al. Burden and Cost of Gastrointestinal, Liver, and Pancreatic Diseases in the United States: Update 2021. Gastroenterology. 2022;162(2):621-44.

4. Eusebi LH, Ratnakumaran R, Yuan Y, Solaymani-Dodaran M, Bazzoli F, Ford AC. Global prevalence of, and risk factors for, gastro-oesophageal reflux symptoms: a meta-analysis. Gut. 2018;67(3):430-40.

5. Fund WCR, Research AIfC. Food, nutrition, physical activity, and the prevention of cancer: a global perspective: Amer Inst for Cancer Research; 2007.

6. Piovani D, Danese S, Peyrin-Biroulet L, Nikolopoulos GK, Lytras T, Bonovas S. Environmental Risk Factors for Inflammatory Bowel Diseases: An Umbrella Review of Meta-analyses. Gastroenterology. 2019;157(3):647-59 e4.

7. McGee EE, Jackson SS, Petrick JL, Van Dyke AL, Adami HO, Albanes D, et al. Smoking, Alcohol, and Biliary Tract Cancer Risk: A Pooling Project of 26 Prospective Studies. J Natl Cancer Inst. 2019;111(12):1263-78.

8. Yadav D, Whitcomb DC. The role of alcohol and smoking in pancreatitis. Nat Rev

Gastroenterol Hepatol. 2010;7(3):131-45.

9. Bujanda L. The effects of alcohol consumption upon the gastrointestinal tract. Am J Gastroenterol. 2000;95(12):3374-82.

10. Laszkowska M, Rodriguez S, Kim J, Hur C. Heavy Alcohol Use Is Associated With Gastric Cancer: Analysis of the National Health and Nutrition Examination Survey From 1999 to 2010. Am J Gastroenterol. 2021;116(5):1083-6.

11. McNabb S, Harrison TA, Albanes D, Berndt SI, Brenner H, Caan BJ, et al. Meta-analysis of 16 studies of the association of alcohol with colorectal cancer. Int J Cancer. 2020;146(3):861-73.

12. Simpson RF, Hermon C, Liu B, Green J, Reeves GK, Beral V, et al. Alcohol drinking patterns and liver cirrhosis risk: analysis of the prospective UK Million Women Study. Lancet Public Health. 2019;4(1):e41-e8.

13. Roberts W, Verplaetse T, Peltier MKR, Moore KE, Gueorguieva R, McKee SA. Prospective association of e-cigarette and cigarette use with alcohol use in two waves of the Population Assessment of Tobacco and Health. Addiction. 2020;115(8):1571-9.

14. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet. 2019;51(2):237-44.

15. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89-98.

16. Yuan S, Larsson SC. Adiposity, diabetes, lifestyle factors and risk of gastroesophageal

509 reflux disease: a Mendelian randomization study. Eur J Epidemiol. 2022.

510 17. Larsson SC, Carter P, Kar S, Vithayathil M, Mason AM, Michaëlsson K, et al. Smoking,

511 alcohol consumption, and cancer: A mendelian randomisation study in UK Biobank and

512 international genetic consortia participants. PLoS Med. 2020;17(7):e1003178.

513 18. Yuan S, Larsson SC. Genetically Predicted Adiposity, Diabetes, and Lifestyle Factors in

514 Relation to Diverticular Disease. Clin Gastroenterol Hepatol. 2022;20(5):1077-84.

515 19. Yuan S, Chen J, Li X, Fan R, Arsenault B, Gill D, et al. Lifestyle and metabolic factors for

516 nonalcoholic fatty liver disease: Mendelian randomization study. Eur J Epidemiol. 2022.

517 20. Chen L, Yang H, Li H, He C, Yang L, Lv G. Insights into modifiable risk factors of

518 cholelithiasis: A Mendelian randomization study. Hepatology. 2022;75(4):785-96.

519 21. Yuan S, Giovannucci EL, Larsson SC. Gallstone disease, diabetes, calcium, triglycerides,

520 smoking and alcohol consumption and pancreatitis risk: Mendelian randomization study. NPJ

521 Genom Med. 2021;6(1):27.

522 22. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open

523 access resource for identifying the causes of a wide range of complex diseases of middle and

524 old age. PLoS Med. 2015;12(3):e1001779.

525 23. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner K, et al. FinnGen:

526 Unique genetic insights from combining isolated population and national health register data.

527 medRxiv. 2022:2022.03.03.22271360.

528 24. Wootton RE, Richmond RC, Stuijfzand BG, Lawn RB, Sallis HM, Taylor GMJ, et al.

529 Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a

Mendelian randomisation study. Psychol Med. 2020;50(14):2435-43.

25. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015;47(9):979-86.

26. Guindo-Martinez M, Amela R, Bonas-Guarch S, Puiggros M, Salvoro C, Miguel-Escalada I, et al. The impact of non-additive genetic associations on age-related complex diseases. Nat Commun. 2021;12(1):2436.

27. Zhou W, Zhao Z, Nielsen JB, Fritsche LG, LeFaive J, Gagliano Taliun SA, et al. Scalable generalized linear mixed model for region-based association tests in large biobanks and cohorts. Nat Genet. 2020;52(6):634-9.

28. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017;46(6):1734-9.

29. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017;32(5):377-89.

30. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018;50(5):693-8.

31. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol. 2013;42(5):1497-501.

32. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base

platform supports systematic causal inference across the human phenome. Elife. 2018;7.

33. Castro C, Peleteiro B, Lunet N. Modifiable factors and esophageal cancer: a systematic review of published meta-analyses. J Gastroenterol. 2018;53(1):37-51.

34. Kato I, Nomura AM, Stemmermann GN, Chyou PH. A prospective study of gastric and duodenal ulcer and its relation to smoking, alcohol, and diet. Am J Epidemiol. 1992;135(5):521-30.

35. Nordenstedt H, Graham DY, Kramer JR, Rugge M, Verstovsek G, Fitzgerald S, et al. Helicobacter pylori-negative gastritis: prevalence and risk factors. Am J Gastroenterol. 2013;108(1):65-71.

36. Zhang YB, Pan XF, Chen J, Cao A, Zhang YG, Xia L, et al. Combined lifestyle factors, incident cancer, and cancer mortality: a systematic review and meta-analysis of prospective cohort studies. Br J Cancer. 2020;122(7):1085-93.

37. Talley NJ, Powell N, Walker MM, Jones MP, Ronkainen J, Forsberg A, et al. Role of smoking in functional dyspepsia and irritable bowel syndrome: three random population-based studies. Aliment Pharmacol Ther. 2021;54(1):32-42.

38. Aune D, Sen A, Leitzmann MF, Tonstad S, Norat T, Vatten LJ. Tobacco smoking and the risk of diverticular disease - a systematic review and meta-analysis of prospective studies. Colorectal Dis. 2017;19(7):621-33.

39. Liu B, Balkwill A, Roddam A, Brown A, Beral V, Million Women Study C. Separate and joint effects of alcohol and smoking on the risks of cirrhosis and gallbladder disease in middle-aged women. Am J Epidemiol. 2009;169(2):153-60.

572 40. Aune D, Vatten LJ, Boffetta P. Tobacco smoking and the risk of gallbladder disease. Eur J
573 Epidemiol. 2016;31(7):643-53.
574 41. Aune D, Mahamat-Saleh Y, Norat T, Riboli E. Tobacco smoking and the risk of
575 pancreatitis: A systematic review and meta-analysis of prospective studies. Pancreatology.
576 2019;19(8):1009-22.
577 42. Montgomery SM, Pounder RE, Wakefield AJ. Smoking in adults and passive smoking in
578 children are associated with acute appendicitis. Lancet. 1999;353(9150):379.
579 43. Park S, Chun J, Han KD, Soh H, Kang EA, Lee HJ, et al. Dose-response relationship
580 between cigarette smoking and risk of ulcerative colitis: a nationwide population-based study.
581 J Gastroenterol. 2019;54(10):881-90.
582 44. Georgiou AN, Ntritsos G, Papadimitriou N, Dimou N, Evangelou E. Cigarette Smoking,
583 Coffee Consumption, Alcohol Intake, and Risk of Crohn's Disease and Ulcerative Colitis: A
584 Mendelian Randomization Study. Inflamm Bowel Dis. 2021;27(2):162-8.
585 45. Botteri E, Borroni E, Sloan EK, Bagnardi V, Bosetti C, Peveri G, et al. Smoking and
586 Colorectal Cancer Risk, Overall and by Molecular Subtypes: A Meta-Analysis. Am J
587 Gastroenterol. 2020;115(12):1940-9.
588 46. Dimou N, Yarmolinsky J, Bouras E, Tsilidis KK, Martin RM, Lewis SJ, et al. Causal Effects
589 of Lifetime Smoking on Breast and Colorectal Cancer Risk: Mendelian Randomization Study.
590 Cancer Epidemiol Biomarkers Prev. 2021;30(5):953-64.
591 47. Mizrahi JD, Surana R, Valle JW, Shroff RT. Pancreatic cancer. Lancet.
592 2020;395(10242):2008-20.

48. Yu C, Tang H, Guo Y, Bian Z, Yang L, Chen Y, et al. Hot Tea Consumption and Its Interactions With Alcohol and Tobacco Use on the Risk for Esophageal Cancer: A Population-Based Cohort Study. Ann Intern Med. 2018;168(7):489-97.

49. Roerecke M, Vafaei A, Hasan OSM, Chrystoja BR, Cruz M, Lee R, et al. Alcohol Consumption and Risk of Liver Cirrhosis: A Systematic Review and Meta-Analysis. Am J Gastroenterol. 2019;114(10):1574-86.

50. Samokhvalov AV, Rehm J, Roerecke M. Alcohol Consumption as a Risk Factor for Acute and Chronic Pancreatitis: A Systematic Review and a Series of Meta-analyses. EBioMedicine. 2015;2(12):1996-2002.

51. Ryan-Harshman M, Aldoori W. How diet and lifestyle affect duodenal ulcers. Review of the evidence. Can Fam Physician. 2004;50:727-32.

52. Bagnardi V, Rota M, Botteri E, Tramacere I, Islami F, Fedirko V, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. Br J Cancer. 2015;112(3):580-93.

53. Klatsky AL. Diet, alcohol, and health: a story of connections, confounders, and cofactors. Am J Clin Nutr. 2001;74(3):279-80.

54. Arnson Y, Shoenfeld Y, Amital H. Effects of tobacco smoke on immunity, inflammation and autoimmunity. J Autoimmun. 2010;34(3):J258-65.

55. Givi ME, Folkerts G, Wagenaar GT, Redegeld FA, Mortaz E. Cigarette smoke differentially modulates dendritic cell maturation and function in time. Respir Res. 2015;16:131.

614 56. Noakes PS, Hale J, Thomas R, Lane C, Devadason SG, Prescott SL. Maternal smoking
615 is associated with impaired neonatal toll-like-receptor-mediated immune responses. Eur
616 Respir J. 2006;28(4):721-9.
617 57. Mandrekar P, Szabo G. Signalling pathways in alcohol-induced liver inflammation. J
618 Hepatol. 2009;50(6):1258-66.
619 58. Bishehsari F, Magno E, Swanson G, Desai V, Voigt RM, Forsyth CB, et al. Alcohol and
620 Gut-Derived Inflammation. Alcohol Res. 2017;38(2):163-71.
621 59. Leone FT, Zhang Y, Evers-Casey S, Evins AE, Eakin MN, Fathi J, et al. Initiating
622 Pharmacologic Treatment in Tobacco-Dependent Adults. An Official American Thoracic
623 Society Clinical Practice Guideline. Am J Respir Crit Care Med. 2020;202(2):e5-e31.
624 60. Camilleri M, Malhi H, Acosta A. Gastrointestinal Complications of Obesity.
625 Gastroenterology. 2017;152(7):1656-70.
626 61. Krishnan B, Babu S, Walker J, Walker AB, Pappachan JM. Gastrointestinal complications
627 of diabetes mellitus. World J Diabetes. 2013;4(3):51-63.
628 62. Abnet CC, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma.
629 Gastroenterology. 2018;154(2):360-73.
630 63. Coleman HG, Xie SH, Lagergren J. The Epidemiology of Esophageal Adenocarcinoma.
631 Gastroenterology. 2018;154(2):390-405.
632 64. Jorgenson E, Thai KK, Hoffmann TJ, Sakoda LC, Kvale MN, Banda Y, et al. Genetic
633 contributors to variation in alcohol consumption vary by race/ethnicity in a large multi-ethnic
634 genome-wide association study. Mol Psychiatry. 2017;22(9):1359-67.

635 65. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample
636 Mendelian randomization. Genet Epidemiol. 2016;40(7):597-608.
637

Table and figure legends

Table 1. Associations of genetic predisposition to smoking initiation with 24 gastrointestinal diseases in univariable and multivariable Mendelian randomization analyses.

Table 2. Associations of genetically predicted alcohol consumption with 24 gastrointestinal diseases in univariable and multivariable Mendelian randomization analyses.

Fig 1. Overview of the present study design. GERA, Genetic Epidemiology Research on Aging; IIBDGC, the International Inflammatory Bowel Disease Genetics Consortium; MR, Mendelian randomization; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier; SNP, single nucleotide polymorphism.

Fig 2. Summary of associations of genetically predicted smoking initiation, lifetime smoking, and alcohol consumption with 24 gastrointestinal diseases. UVMR, univariable Mendelian randomization; MVMR, multivariable Mendelian randomization. The numbers in the box are the odds ratios for associations of exposure for gastrointestinal diseases. The association with a *P*-value <0.05 but Benjamini-Hochberg adjusted *P*-value >0.05 was regarded suggestive and the association with a Benjamini–Hochberg adjusted P-value <0.05 were deemed significant.

Table 1. Associations of genetic predisposition to smoking initiation with 24 gastrointestinal diseases in univariable and multivariable Mendelian randomization analyses.

Disease		Total cases	Total controls	UVMR			MVMR adjusted for alcohol consumption	
				OR (95% CI)	P value	I^2 (95% CI)	OR (95% CI)	P value
Upper Gastrointestinal Diseases	Gastroesophageal reflux	34,135	634,629	1.28 (1.20, 1.37)	$4.09\times10^{-14*}$	46.24	1.65 (1.35, 2.02)	$1.38\times10^{-6*}$
	Esophageal cancer	1,130	702,116	1.67 (1.24, 2.25)	$6.84\times10^{-4*}$	22.68	4.78 (2.10, 10.90)	$1.97\times10^{-4*}$
	Gastric ulcer	8,651	666,879	1.54 (1.37, 1.72)	$3.83\times10^{-14*}$	44.96	1.95 (1.40, 2.71)	$7.31\times10^{-5*}$
	Duodenal ulcer	5,713	666,879	1.53 (1.34, 1.75)	$8.47\times10^{-10*}$	0.00	1.64 (1.07, 2.52)	0.024
	Acute Gastritis	3,048	643,478	1.29 (1.09, 1.53)	0.003^{*}	0.00	1.54 (0.91, 2.62)	0.106
	Chronic gastritis	7,975	643,478	1.33 (1.18, 1.49)	$1.55\times10^{-6*}$	77.04	1.33 (0.96, 1.86)	0.091
	Gastric cancer	1,608	701,472	1.42 (1.13, 1.79)	0.002^{*}	0.00	2.29 (1.14, 4.59)	0.020
Lower Gastrointestinal Diseases	Irritable bowel disease	15,718	641,489	1.22 (1.12, 1.32)	$3.50\times10^{-6*}$	11.84	1.43 (1.10, 1.85)	0.008*
	Celiac disease	4,808	631,700	0.82 (0.66, 1.02)	0.071	0.00	0.87 (0.53, 1.43)	0.59
	Diverticular disease	50,065	587,969	1.25 (1.18, 1.33)	$5.23\times10^{-14*}$	67.29	1.56 (1.30, 1.87)	$1.41\times10^{-6*}$
	Crohn's disease	10,846	645,718	1.25 (1.11, 1.40)	$3.03\times10^{-4*}$	0.00	1.48 (1.01, 2.16)	0.042
	Ulcerative colitis	16,770	651,255	1.15 (1.04, 1.26)	0.004^{*}	0.00	0.94 (0.71, 1.25)	0.677
	Colorectal cancer	9,519	686,953	1.03 (0.92, 1.14)	0.632	29.94	1.03 (0.76, 1.39)	0.841
Hepatobiliary and Pancreatic Diseases	Non-alcoholic fatty liver disease	3,242	707,631	1.49 (1.26, 1.76)	$3.82\times10^{-6*}$	0.00	2.11 (1.15, 3.88)	0.016^{*}
	Alcoholic liver disease	2,955	680,369	1.99 (1.65, 2.41)	$1.49\times10^{-12*}$	92.68	2.26 (1.26, 4.03)	0.006
	Cirrhosis	5,904	706,200	1.68 (1.40, 2.02)	$3.39\times10^{-8*}$	0.00	1.92 (1.06, 3.47)	0.032
	Liver cancer	714	702,008	1.57 (1.13, 2.17)	0.007^{*}	0.00	1.96 (0.73, 5.25)	0.183
	Cholangitis	1,708	664,749	1.02 (0.80, 1.29)	0.892	0.00	1.31 (0.61, 2.84)	0.489
	Cholecystitis	5,893	664,749	1.47 (1.29, 1.68)	$4.71\times10^{-9*}$	84.72	2.38 (1.57, 3.60)	$4.14\times10^{-5*}$
	Cholelithiasis	42,510	664,749	1.20 (1.13, 1.27)	$5.75\times10^{-9*}$	0.00	1.33 (1.02, 1.73)	0.035
	Acute pancreatitis	6,634	679,713	1.39 (1.23, 1.56)	$6.71\times10^{-8}*$	79.71	1.55 (1.04, 2.31)	0.031
	Chronic pancreatitis	3,173	679,713	1.38 (1.17, 1.64)	$1.79\times10^{-4}*$	0.00	1.27 (0.74, 2.16)	0.384
	Pancreatic cancer	1,643	701,472	1.00 (0.79, 1.26)	0.999	67.21	2.08 (1.06, 4.10)	0.034
Other	Acute appendicitis	25,361	690,149	1.15 (1.08, 1.23)	$1.27\times10^{-5}*$	0.00	1.15 (0.92, 1.44)	0.221

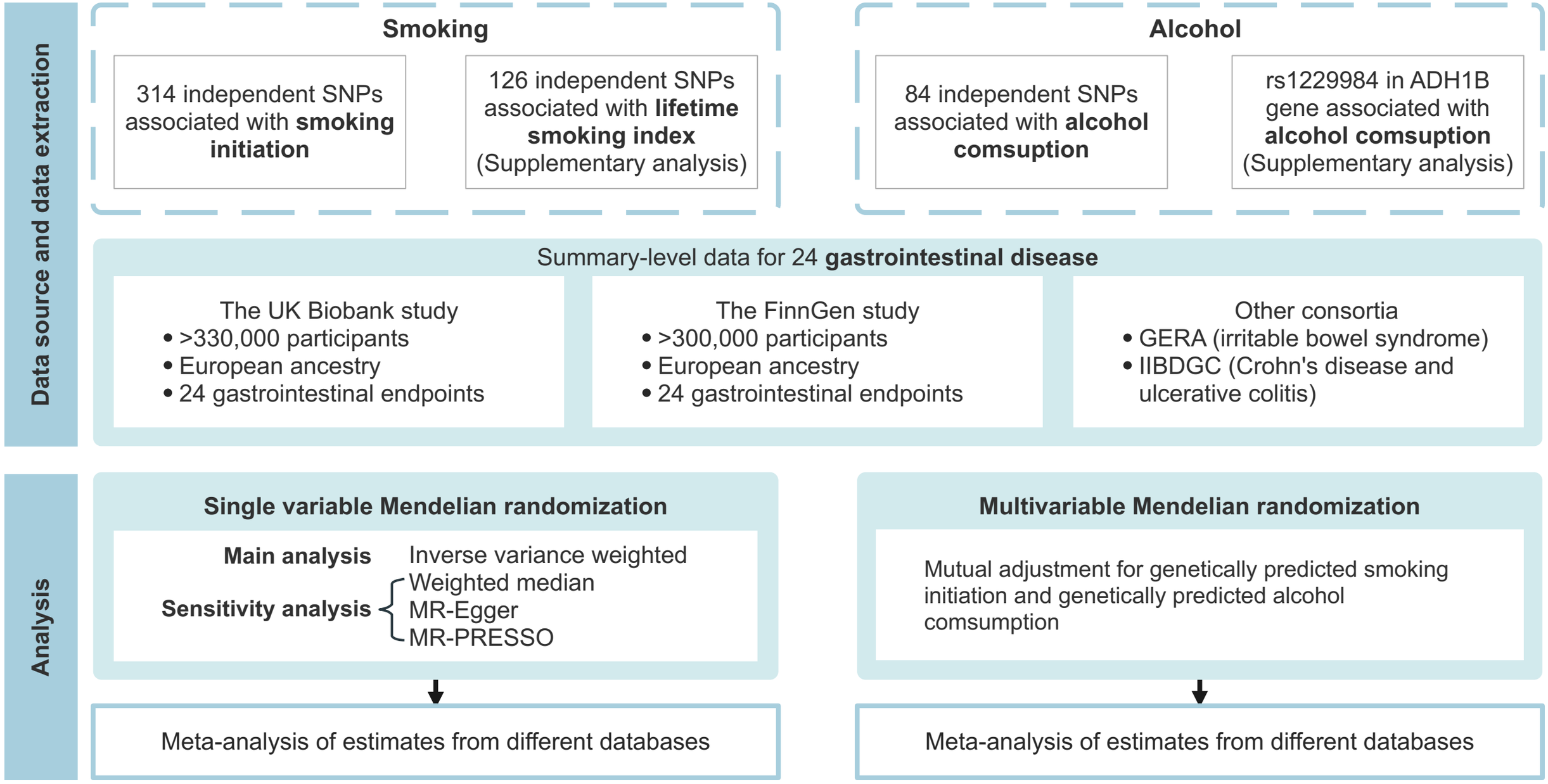
662 UVMR, univariable Mendelian randomization; MVMR, multivariable Mendelian randomization; OR, odds ratio; CI, confidence interval. *Significant association
663 after multiple testing.
664
665

Table 2. Associations of genetically predicted alcohol consumption with 24 gastrointestinal diseases in univariable and multivariable Mendelian randomization analyses.

Disease		Total cases	Total controls	UVMR			MVMR adjusted for smoking initiation	
				OR (95% CI)	P value	I^2 (95% CI)	OR (95% CI)	P value
Upper Gastrointestinal Diseases	Gastroesophageal reflux	34,135	634,629	0.99 (0.81, 1.21)	0.893	46.24	0.88 (0.72, 1.08)	0.219
	Esophageal cancer	1,130	702,116	2.86 (1.18, 6.91)	0.020	22.68	1.28 (0.59, 2.82)	0.533
	Gastric ulcer	8,651	666,879	1.30 (0.95, 1.77)	0.098	44.96	1.06 (0.77, 1.47)	0.721
	Duodenal ulcer	5,713	666,879	1.92 (1.23, 3.00)	0.004*	0.00	1.54 (1.01, 2.34)	0.045
	Acute Gastritis	3,048	643,478	0.99 (0.58, 1.69)	0.960	0.00	0.88 (0.52, 1.48)	0.621
	Chronic gastritis	7,975	643,478	1.33 (0.90, 1.95)	0.147	77.04	1.33 (0.93, 1.89)	0.115
	Gastric cancer	1,608	701,472	1.57 (0.75, 3.30)	0.233	0.00	1.59 (0.79, 3.21)	0.194
Lower Gastrointestinal Diseases	Irritable bowel disease	15,718	641,489	0.73 (0.57, 0.93)	0.012	11.84	0.74 (0.57, 0.97)	0.027
	Celiac disease	4,808	631,700	0.87 (0.53, 1.42)	0.097	0.00	1.04 (0.64, 1.68)	0.887
	Diverticular disease	50,065	587,969	0.95 (0.79, 1.13)	0.553	67.29	0.94 (0.79, 1.13)	0.527
	Crohn's disease	10,846	645,718	0.91 (0.62, 1.32)	0.613	0.00	0.74 (0.53, 1.05)	0.088
	Ulcerative colitis	16,770	651,255	1.11 (0.82, 1.50)	0.509	0.00	0.88 (0.67, 1.15)	0.358
	Colorectal cancer	9,519	686,953	1.09 (0.76, 1.55)	0.649	29.94	1.28 (0.95, 1.72)	0.098
Hepatobiliary and Pancreatic Diseases	Non-alcoholic fatty liver disease	3,242	707,631	1.20 (0.63, 2.28)	0.574	0.00	0.99 (0.54, 1.79)	0.962
	Alcoholic liver disease	2,955	680,369	14.35 (7.69, 26.81)	6.32×10^{-17}*	92.68	9.60 (5.28, 17.46)	1.25×10^{-13}*
	Cirrhosis	5,904	706,200	2.96 (1.50, 5.85)	0.002*	0.00	2.41 (1.29, 4.52)	0.006*
	Liver cancer	714	702,008	1.16 (0.43, 3.11)	0.775	0.00	0.76 (0.29, 2.02)	0.585
	Cholangitis	1,708	664,749	0.96 (0.44, 2.08)	0.912	0.00	0.72 (0.33, 1.55)	0.397
	Cholecystitis	5,893	664,749	1.36 (0.91, 2.03)	0.132	84.72	0.96 (0.64, 1.45)	0.862
	Cholelithiasis	42,510	664,749	1.02 (0.75, 1.39)	0.878	0.00	1.03 (0.79, 1.35)	0.801
	Acute pancreatitis	6,634	679,713	1.36 (0.91, 2.03)	0.128	79.71	1.17 (0.78, 1.75)	0.456
	Chronic pancreatitis	3,173	679,713	2.96 (1.80, 4.89)	2.13×10^{-5}*	0.00	3.24 (1.86, 5.64)	3.18×10^{-5}*
	Pancreatic cancer	1,643	701,472	0.63 (0.32, 1.26)	0.193	67.21	0.79 (0.40, 1.56)	0.496
Other	Acute appendicitis	25,361	690,149	0.80 (0.63, 1.01)	0.063	0.00	0.77 (0.61, 0.97)	0.024

UVMR, univariable Mendelian randomization; MVMR, multivariable Mendelian randomization; OR, odds ratio; CI, confidence interval.

669 *Significant association after multiple testing
670


Data source and data extraction
Smoking
314 independent SNPs associated with smoking initiation
126 independent SNPs associated with lifetime smoking index (Supplementary analysis)
Alcohol
84 independent SNPs associated with alcohol comsuption
rs1229984 in ADH1B gene associated with alcohol comsuption (Supplementary analysis)
Summary-level data for 24 gastrointestinal disease
The UK Biobank study
• >330,000 participants
• European ancestry
• 24 gastrointestinal endpoints
The FinnGen study
• >300,000 participants
• European ancestry
• 24 gastrointestinal endpoints
Other consortia
• GERA (irritable bowel syndrome)
• IIBDGC (Crohn's disease and ulcerative colitis)
Analysis
Single variable Mendelian randomization
Main analysis
Inverse variance weighted
Sensitivity analysis
Weighted median
MR-Egger
MR-PRESSO
Multivariable Mendelian randomization
Mutual adjustment for genetically predicted smoking initiation and genetically predicted alcohol comsumption
Meta-analysis of estimates from different databases
Meta-analysis of estimates from different databases

	Smoking initiation UVMR	Smoking initiation MVMR	Lifetime smoking index	Alcohol consumption UVMR	Alcohol consumption MVMR	Alcohol consumption instrumented by rs1229984
Gastroesophageal reflux	1.28	1.65	1.42	0.99	0.88	0.87
Esophageal cancer	1.67	4.78	2.34	2.86	1.28	8.90
Gastric ulcer	1.54	1.95	1.87	1.30	1.06	1.34
Duodenal ulcer	1.53	1.64	2.46	1.92	1.54	1.79
Acute gastritis	1.29	1.54	1.40	0.99	0.88	0.52
Chronic gastritis	1.33	1.33	1.27	1.33	1.33	2.71
Gastric cancer	1.42	2.29	1.32	1.57	1.59	1.89
Irritable bowel syndrome	1.22	1.43	1.29	0.73	0.74	1.06
Celiac disease	0.82	0.87	1.04	0.87	1.04	1.60
Diverticular disease	1.25	1.56	1.30	0.95	0.94	0.98
Crohn’s disease	1.25	1.48	1.07	0.91	0.74	0.94
Ulcerative colitis	1.15	0.94	0.99	1.11	0.88	1.06
Colorectal cancer	1.03	1.03	1.13	1.09	1.28	0.87
Non−alcoholic fatty liver disease	1.49	2.11	2.38	1.20	0.99	6.85
Alcoholic liver disease	1.99	2.26	2.89	14.35	9.60	172.97
Cirrhosis	1.68	1.92	2.10	2.96	2.41	6.79
Liver cancer	1.57	1.96	2.47	1.16	0.76	4.60
Cholangitis	1.02	1.31	1.24	0.96	0.72	0.73
Cholecystitis	1.47	2.38	1.79	1.36	0.96	1.19
Cholelithiasis	1.20	1.33	1.41	1.02	1.03	0.69
Acute pancreatitis	1.39	1.55	1.64	1.36	1.17	0.75
Chronic pancreatitis	1.38	1.27	2.35	2.96	3.24	4.91
Pancreatic cancer	1.00	2.08	2.09	0.63	0.79	1.19
Acute appendicitis	1.15	1.15	1.17	0.80	0.77	0.31

Significant positive association
Suggestive positive association
No significant association
Suggestive inverse association
Significant inverse association

Smoking, Alcohol Consumption, and 24 Gastrointestinal Diseases: Mendelian Randomization Analysis

Running head: Smoking, alcohol intake, and 24 gastrointestinal diseases

Shuai Yuan[1,2 #], Jie Chen[1,3 #], Xixian Ruan[3 #], Yuhao Sun[1], Ke Zhang[4,5], Xiaoyan Wang[3*], Xue Li[1,6*], Dipender Gill[7], Stephen Burgess[8,9], Edward L Giovannucci[10,11], Susanna C. Larsson[2,12]

Affiliations

[1] School of Public Health and The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China.

[2] Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

[3] Department of Gastroenterology, The Third Xiangya Hospital, Central South University, Changsha, China.

[4] Key Laboratory of Growth Regulation and Translational Research of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, China.

[5] Westlake Intelligent Biomarker Discovery Lab, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, China.

[6] Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh, UK.

[7] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.

[8] MRC Biostatistics Unit, University of Cambridge, Cambridge, UK.

[9] Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

[10] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[11] Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[12] Unit of Medical Epidemiology, Department of Surgical Sciences, Uppsala University, Uppsala, Sweden.

[#] Shuai Yuan, Jie Chen, and Xixian Ruan share first authorship

[*]Xiaoyan Wang and Xue Li share corresponding authorship

Correspondence to

Xue Li, School of Public Health and the Second Affiliated Hospital, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China. E-mail: xue.li@ed.ac.uk

Xiaoyan Wang, Department of Gastroenterology, The Third Xiangya Hospital, Central South University, 138 Tongzipo Road, Changsha, Hunan 410013, P.R. China. E-mail: wangxiaoyan@csu.edu.cn

Word count: abstract, 252; text, 3801

Number of tables and figures: 2 tables and 2 figures

Abstract

Background: Whether the positive associations of smoking and alcohol consumption with gastrointestinal diseases are causal is uncertain. We conducted this Mendelian randomization (MR) to comprehensively examine associations of smoking and alcohol consumption with common gastrointestinal diseases.

Methods: Genetic variants associated with smoking initiation and alcohol consumption at the genome-wide significance level were selected as instrumental variables. Genetic associations with 24 gastrointestinal diseases were obtained from the UK Biobank, FinnGen study, and other large consortia. Univariable and multivariable MR analyses were conducted to estimate the overall and independent MR associations after mutual adjustment for genetic liability to smoking and alcohol consumption.

Results: Genetic predisposition to smoking initiation was associated with increased risk of 20 of 24 gastrointestinal diseases, including 7 upper gastrointestinal diseases (gastroesophageal reflux, esophageal cancer, gastric ulcer, duodenal ulcer, acute gastritis, chronic gastritis and gastric cancer), 4 lower gastrointestinal diseases (irritable bowel syndrome, diverticular disease, Crohn's disease and ulcerative colitis), 8 hepatobiliary and pancreatic diseases (non-alcoholic fatty liver disease, alcoholic liver disease, cirrhosis, liver cancer, cholecystitis, cholelithiasis, acute and chronic pancreatitis), and acute appendicitis. Fifteen out of 21 associations persisted after adjusting for genetically-predicted alcohol consumption. Genetically-predicted higher alcohol consumption was associated with increased risk of duodenal cancer, alcoholic liver disease, cirrhosis, and chronic pancreatitis; however, the association for duodenal ulcer did not remain after adjustment for genetic predisposition to smoking initiation.

Conclusion: This study provides MR evidence supporting causal associations of smoking with a broad range of gastrointestinal diseases, whereas alcohol consumption was associated with only a few gastrointestinal diseases.

Funding: The Natural Science Fund for Distinguished Young Scholars of Zhejiang Province; National Natural Science Foundation of China; Key Project of Research and Development Plan of Hunan Province; the Swedish Heart Lung Foundation; the Swedish Research Council; the Swedish Cancer Society.

Keywords: alcohol consumption; gastrointestinal diseases; Mendelian randomization; smoking

Introduction

Tobacco smoking and alcohol consumption are leading causes of the global burden of disease and are major contributors to premature mortality (1, 2). Gastrointestinal diseases account for considerable health care use and expenditures, and a holistic approach to lifestyle interventions may result in more health gains and less economic burdens (3). Population-based studies have identified tobacco smoking as a risk factor for several gastrointestinal diseases, including gastroesophageal reflux disease (4), esophageal cancer (5), Crohn's disease (6), liver cancer (7), and pancreatitis (8). Evidence on the association between tobacco smoking and risk of other gastrointestinal diseases is limited and inconsistent. Like smoking, heavy alcohol consumption has been associated with increased risk of several gastrointestinal outcomes, including gastritis (9), gastric cancer (10), colorectal cancer (11), cirrhosis (12), liver cancer (7), and pancreatitis (8). However, whether these associations are all causal remain unestablished since most of the evidence was obtained from observational studies where the results may be biased by reverse causality and confounding. Of note, even though reverse causality may not be an issue in the studies for any of studied gastroenterological outcomes, it might exist for certain gastroenterological diseases causing pain, which smoker patients may try to increase smoking dose to mitigate via an intake of higher levels of nicotine. In addition, as smoking and alcohol consumption are phenotypically and genetically correlated (13, 14), the independent impacts of smoking and alcohol consumption on

gastrointestinal diseases are unclear. Establishing the causal association of tobacco smoking and alcohol consumption with gastrointestinal diseases is crucial, as this provides further evidence for subsequent recommending public policies and clinical interventions.

Mendelian randomization (MR) is an epidemiological approach that utilizes genetic variants as an instrument to strengthen the causal inference in an exposure-outcome association (15). MR is by nature not prone to confounding since genetic variants are randomly assorted at conception and thus unrelated to environmental and self-adopted factors that usually act as confounders. Additionally, this method can minimize reverse causality since fixed alleles are unaffected by the onset and progression of disease. Previous MR studies have examined the associations of smoking and alcohol consumption with several gastrointestinal diseases (16-21). Nevertheless, whether smoking and alcohol consumption exert influence on a wide range of gastrointestinal outcomes have not been investigated in a comprehensive way. A thorough investigation on the gastrointestinal consequences of smoking and alcohol drinking is of great importance to develop non-pharmacological interventions on gastrointestinal diseases. Here, we conducted an MR investigation of the associations of smoking and alcohol consumption with the risk of common gastrointestinal diseases to fill in above knowledge gaps.

Materials and Methods

Fig 1 shows the study design overview. The study was based on publicly available genome-wide association studies (GWAS), and the detailed information on used studies was presented in **S1 Table**. The genetic associations were estimated using data from the UK Biobank study (22), the FinnGen study (23), and several large consortia. The summary effect estimates were combined using meta-analysis for each gastrointestinal disease from different data resources. Included studies had been approved by corresponding institutional review boards and ethical committees, and consent forms had been signed by all participants.

Instrumental variable selection

A total of 378 and 99 single nucleotide polymorphisms (SNPs) associated with smoking initiation (a binary phenotype indicating whether an individual had ever being a regular smoker, 1,232,091 individuals of European descent) and alcohol consumption (log-transformed drinks per week, 941,280 individuals of European descent) at the genome-wide significance threshold ($P<5\times10^{-8}$) were identified by the GWAS and Sequencing Consortium of Alcohol and Nicotine use (GSCAN) study (14). These SNPs explained approximately 2.3% and 0.3% of the variation in smoking initiation and alcohol consumption, respectively (14). SNPs in linkage disequilibrium (defined as $r^2>0.01$ or clump distance <10,000 kb) and with the weaker associations with the exposure were removed, leaving 314 independent SNPs as

141 instrumental variables for smoking initiation and 84 for alcohol consumption. Smoking
142 initiation and alcohol consumption shared two index genetic variants, which were
143 rs1713676 and rs11692435. Considering partial sample overlap (around 30%)
144 between the GSCAN study with full data and the UK Biobank study (14), we
145 performed sensitivity analyses for smoking initiation and alcohol consumption using
146 summary statistics data from the analysis excluding the UK Biobank and 23andMe.
147 For a supplementary analysis of smoking behavior, we used 126 SNPs associated
148 with a lifetime smoking index that considered smoking duration, heaviness, and
149 cessation (24). The set of genetic instruments captured around 0.36% of the variance
150 in lifetime smoking (24). We also conducted a sensitivity analysis using rs1229984 in
151 *ADH1B* gene that encodes alcohol dehydrogenase 1B enzyme as the genetic
152 instrument for alcohol consumption to minimize pleiotropy. Detailed information on
153 used SNPs is presented in **S2 Table**.
154
155 **Gastrointestinal disease data sources**
156 Genetic associations with 24 gastrointestinal diseases were obtained from the UK
157 Biobank study (22), the FinnGen study (23), and two large consortia, including the
158 International Inflammatory Bowel Disease Genetics Consortium (IIBDGC) (25) and
159 Genetic Epidemiology Research on Aging (GERA) (26). Included outcomes were
160 classified into four major categories according to the disease onset site: 1) upper
161 gastrointestinal diseases (gastroesophageal reflux disease, esophageal cancer,

gastric ulcer, acute gastritis, chronic gastritis, and gastric cancer); 2) lower gastrointestinal diseases (irritable bowel disease, celiac disease, diverticular disease, Crohn's disease, ulcerative colitis, and colorectal cancer); 3) hepatobiliary and pancreatic disease (non-alcoholic fatty liver disease, alcoholic liver disease, cirrhosis, liver cancer, cholangitis, cholecystitis, cholelithiasis, acute pancreatitis, chronic pancreatitis, and pancreatic cancer); 4) other (acute appendicitis).

The UK Biobank study is a large multicenter cohort study of 500,000 participants recruited in the United Kingdom between 2006 and 2010 (22). We used the summary statistics of European ancestry from GWAS conducted by Lee lab where the gastrointestinal outcomes were defined by codes of the International Classification of Diseases 9th Revision (ICD-9) and ICD-10 (27). Genetic associations were estimated by logistic regression with adjustment for sex, birth year, and the first four genetic principal components. For the FinnGen study, we used summary-level data on the genetic associations with gastrointestinal diseases from the last publicly available R7 data release (23). The FinnGen study is a nationwide genetic study where genetic and electronic health record data were collected. The gastrointestinal diseases were ascertained by the codes of the ICD-8, ICD-9, and ICD-10. Genome-wide association analyses were adjusted for sex, age, genetic components, and genotyping batch. Summary-level genetic data on Crohn's disease (5,956 cases and 14,927 controls) and ulcerative colitis (6,968 cases and 20,464 controls) were additionally obtained

183 from the IIBDGC (25) and data on irritable bowel syndrome (3,117 cases and 53,520
184 controls) were obtained from the GERA (26). Detailed diagnostic codes are listed in
185 **S3 Table.**
186
187 **Statistical analysis**
188 Data were harmonized to omit ambiguous SNPs with non-concordant alleles and
189 palindromic SNPs with ambiguous minor allele frequency (>0.42 and <0.58) were
190 removed from the analysis. The primary MR analyses were performed by the
191 multiplicative random-effects inverse-variance weighted (IVW) method, which
192 provides the most precise estimates though assuming that all SNPs are valid
193 instruments. The analysis of rs1229984 for alcohol consumption was conducted by
194 the Wald method. Estimates for each association from different sources were
195 combined using fixed-effects meta-analysis and the heterogeneity of the associations
196 from different data sources were evaluated by the I^2 statistic. Heterogeneity among
197 SNPs' estimates in each association was assessed by Cochran's Q value.
198 Multivariable MR analyses were conducted to mutually adjust for smoking initiation
199 and alcohol consumption. To detect potential unbalanced pleiotropy (horizontal
200 pleiotropy) and examine the consistency of the associations, three sensitivity
201 analyses including the weighted median (28), MR-Egger (29), and Mendelian
202 randomization pleiotropy residual sum and outlier (MR-PRESSO) (30) analyses were
203 performed. The weighted median method can provide consistent estimates when

more than 50% of the weight comes from valid instrument variants (28). The MR-Egger intercept test can detect unmeasured pleiotropy and MR-Egger regression can generate estimates after accounting for horizontal pleiotropy albeit with less precision (29). The MR-PRESSO method can identify SNP outliers and provide results identical to that from IVW after removal of outliers (30). The *F*-statistic was estimated to quantify instrument strength and an *F*-statistic >10 suggested a sufficiently strong instrument. Power analysis was performed using an online tool (31). The Benjamini-Hochberg correction that controls the false discovery rate was applied to correct for multiple testing. The association with a nominal *P*-value <0.05 but Benjamini-Hochberg adjusted *P*-value >0.05 was regarded suggestive and the association with a Benjamini–Hochberg adjusted *P*-value <0.05 were deemed significant. All analyses were two-sided and performed using the TwoSampleMR (32), MendelianRandomization (28), and MRPRESSO R packages (30) in R software 4.1.2.

Results

The *F*-statistic for each genetic variant was above 10, suggesting a good strength of used genetic instruments (**S2 Table**). Most associations were well powered (**S4 Table**). For smoking initiation, there was 80% power to detect the smallest odds ratio (OR) ranging from 1.08 to 1.40 for included outcomes. Although power was lower for

alcohol consumption, it was adequate to detect a moderate effect size for most common gastrointestinal diseases.

Smoking and gastrointestinal diseases

Genetic predisposition to smoking initiation was associated with 20 of the 24 studied gastrointestinal diseases and all these associations remained after multiple comparison correction (**Table 1 and S5 Table**). In detail, genetic liability to smoking initiation was positively associated with 7 upper gastrointestinal diseases: gastroesophageal reflux (OR, 1.28; 95% confidence interval [CI], 1.20-1.37; $P = 4.09\times10^{-14}$), esophageal cancer (OR, 1.67; 95% CI, 1.24-2.25; $P = 6.84\times10^{-4}$), gastric ulcer (OR, 1.54; 95% CI, 1.37-1.72; $P = 3.83\times10^{-14}$), duodenal ulcer (OR, 1.53; 95%CI, 1.34-1.75; $P = 8.47\times10^{-10}$), acute gastritis (OR, 1.29; 95%CI, 1.09-1.53; $P = 0.003$), chronic gastritis (OR, 1.33; 95%CI, 1.18-1.49; $P = 1.55\times10^{-6}$) and gastric cancer (OR, 1.42; 95%CI, 1.13-1.79; $P = 0.002$); genetic liability to smoking initiation was positively associated with 4 lower gastrointestinal diseases: irritable bowel syndrome (OR, 1.22; 95%CI, 1.12-1.32; $P = 3.50\times10^{-6}$), diverticular disease (OR, 1.25; 95%CI, 1.18-1.33; $P = 5.23\times10^{-14}$), Crohn's disease (OR, 1.25; 95%CI, 1.11-1.40; $P = 3.03\times10^{-4}$) and ulcerative colitis (OR, 1.15; 95%CI, 1.04-1.26; $P = 0.004$); genetic liability to smoking initiation was positively associated with 8 hepatobiliary and pancreatic diseases: non-alcoholic fatty liver disease (OR, 1.49; 95%CI, 1.26-1.76; $P = 3.82\times10^{-6}$), alcoholic liver disease (OR, 1.99; 95%CI, 1.65-2.41; $P =$

245 1.49×10^{-12}), cirrhosis (OR, 1.68; 95%CI, 1.40-2.02; $P = 3.39\times10^{-8}$), liver cancer (OR,
246 1.57; 95%CI, 1.13-2.17; $P = 0.007$), cholecystitis (OR, 1.47; 95%CI, 1.29-1.68; $P =$
247 4.71×10^{-9}), cholelithiasis (OR, 1.20; 95%CI, 1.13-2.27; $P = 5.75\times10^{-9}$), acute
248 pancreatitis (OR, 1.39; 95%CI, 1.23-1.56; $P = 6.71\times10^{-8}$) and chronic pancreatitis
249 (OR, 1.38; 95% CI, 1.17-1.64; $P = 1.79\times10^{-4}$); genetic liability to smoking initiation
250 was positively associated with acute appendicitis (OR, 1.15; 95% CI, 1.08-1.23; $P =$
251 1.27×10^{-5}). Results were consistent in sensitivity analyses. An indication of horizontal
252 pleiotropy was observed in the analysis of esophageal cancer in the FinnGen study
253 (P for MR-Egger intercept <0.05, **S6 Table**). Although MR-PRESSO detected 1 to 3
254 outliers, the associations persisted and remained significant after removal of these
255 out-lying SNPs (**S6 Table**). When using the genetic variants for smoking initiation
256 based on data without the UK Biobank and 23andMe studies, the associations
257 attenuated slightly albeit remained significant after multiple comparisons (Fig S1 and
258 Table S7). All associations were replicated in the supplementary analysis of the
259 lifetime smoking index (**S7 Table**). After correcting for multiple testing, genetically
260 predicted lifetime smoking index was significantly associated with 17 of 24
261 gastrointestinal diseases, where the patterns of associations were generally similar to
262 the analysis for smoking initiation (**Fig S2 and S7 Table)**. In distinction to the
263 analysis of smoking initiation, genetically predicted lifetime smoking index was not
264 significantly associated with acute gastritis, gastric cancer, Crohn's disease, and

ulcerative colitis, whereas genetically predicted lifetime smoking index was significantly associated with pancreatic cancer (OR, 2.09; 95% CI, 1.30-3.36).

In multivariable MR analysis adjusted for genetically predicted alcohol consumption, the associations between genetically predicted smoking initiation and gastrointestinal diseases were consistent with that from univariable MR analysis (**Table 1 and S8 Table**). However, the associations became stronger with wider CIs, in particular the associations for gastrointestinal reflux, esophageal cancer, gastric ulcer, irritable bowel syndrome, diverticular disease, non-alcoholic fatty liver disease, alcoholic liver disease, and cholecystitis (**Table 1**). In addition, the association for pancreatic cancer became suggestive significant from null.

Alcohol consumption and gastrointestinal diseases

Genetically-predicted alcohol consumption was nominally positively associated with esophageal cancer (OR, 2.86; 95% CI, 1.18-6.91; $P = 0.020$), duodenal ulcer (OR, 1.92; 95% CI, 1.23-3.00; $P = 0.004$), alcoholic liver disease (OR, 14.35; 95% CI, 7.69-26.81; $P = 6.32\times10^{-17}$), cirrhosis (OR, 2.96; 95% CI, 1.50-5.85; $P = 0.002$) and chronic pancreatitis (OR, 2.96; 95% CI, 1.80-4.89; $P = 2.13\times10^{-5}$), and nominally inversely associated with irritable bowel disease (OR, 0.73; 95% CI 0.57-0.93; $P = 0.012$) (**Table 2**). After Benjamini–Hochberg correction, the associations for duodenal cancer, alcoholic liver disease, cirrhosis, and chronic pancreatitis remained (**S5 Table**). Results were consistent in sensitivity analyses, and no horizontal pleiotropy

was detected (**S9 Table**). One outlier was detected in the analysis of duodenal ulcer in the FinnGen study, and the association slightly changed after removal of this outlier (**S9 Table**). Results were consistent in the sensitivity analysis where the genetic associations with alcohol consumption were obtained from the genome-wide association analysis excluding the UK Biobank and 23andMe studies (**Fig S3 and S7 Table**). The associations were directionally consistent albeit with wider CIs in the analysis where alcohol consumption was instrumented by rs1229984 (**S10 Table**). The associations for alcoholic liver disease, cirrhosis, and chronic pancreatitis persisted after adjustment for genetic liability to smoking initiation and multiple testing correction (**Table 2 and S8 Table**).

Discussion

We conducted a comprehensive MR investigation to examine the causal role of smoking and alcohol consumption in 24 gastrointestinal diseases and the result summary of this comprehensive analysis is shown in **Fig 2** and **S11 Table**. We found robust associations between genetic predisposition to smoking and increased risk of 21 gastrointestinal outcomes independent of alcohol consumption, showing an extensive impact on gastrointestinal health. In contrast, genetically predicted alcohol consumption was robustly and predominantly associated with increased risk of liver and pancreatic diseases, including alcoholic liver disease, cirrhosis, and chronic pancreatitis after adjustment for smoking.

Corroborating and extending the previous observational studies, our MR investigation strengthened the evidence that smoking has a detrimental effect on gastrointestinal health and increases the risk of a broad range of gastrointestinal diseases, including gastroesophageal reflux disease (4), esophageal cancer (33), gastric and duodenal ulcer (34), gastritis (35), gastric cancer (36), irritable bowel syndrome (37), diverticular disease (38), Crohn's disease (6), cirrhosis (39), liver cancer (7), cholelithiasis (40), acute and chronic pancreatitis (41), and acute appendicitis (42). In line with previous MR studies, the current MR study also found that smoking was associated with increased risk of gastroesophageal reflux disease (16), esophageal cancer (17), gastric cancer (17), diverticular disease (18) non-alcoholic fatty liver disease (19), cholelithiasis (20), and acute and chronic pancreatitis (21). As for ulcerative colitis, traditional observational studies revealed a decreased risk among current smokers (6, 43); however, a recent MR analysis including 12,366 ulcerative colitis cases did not verify this inverse association in the analysis where smoking initiation was instrumented by 363 SNPs (44). Based on data from 3 independent populations, our study provided genetic evidence that smoking was a causal risk factor for ulcerative colitis in the analysis including 16,770 cases. Observational studies found that smoking was associated with an increased risk of colorectal cancer in a dose-dependent manner (45), whereas the positive association was not observed in an MR analysis (17). The current study was in line with the above MR

study and found no strong association between smoking initiation and colorectal cancer risk. Nevertheless, a previous MR analysis with a 52,775 colorectal cancer cases found that genetic prediction to lifetime smoking index was positively associated with risks of colorectal cancer (46), which might imply that our null finding might be caused by insufficient power due to a relatively small sample size. Smoking has been identified as a well-established risk factor for pancreatic cancer (47). Interestingly, despite a null finding on the association of genetic liability to smoking initiation and pancreatic cancer in univariable MR analysis, the association became stronger and suggestively significant after adjusting for genetically predicted alcohol consumption. This might be explained by an inverse association between moderate alcohol consumption and pancreatic cancer. In addition, an adverse effect of smoking on pancreatic cancer was observed when using a smoking index as genetic instrument for lifetime smoking exposure. Our findings also provide novel evidence on the associations of smoking with the higher risk of cholecystitis and alcoholic liver disease independently of alcohol consumption, which need to be verified.

The pathogenic role of alcohol in alcoholic liver disease is well-established and was confirmed also in our MR analysis. Our MR evidence along with previous observational studies also supported alcohol consumption as a risk factor for esophageal cancer (48), cirrhosis (49), and chronic pancreatitis (50). Noteworthy, the association between alcohol consumption and esophageal cancer became positively

nonsignificant in multivariable MR, which possibly explained by the synergistic effect of alcohol and smoking. However, the association between alcohol consumption and duodenal ulcer has been scarcely studied. A meta-analysis including a small number of studies with relatively small sample sizes indicated that alcohol consumption was not associated with duodenal ulcer (51). This null finding is likely due to insufficient power. Alcohol drinking has been associated with increased risk of gastric, colorectal, and liver cancer as well as acute pancreatitis (52). These associations were not supported by our MR study. A possible explanation for this inconsistent findings is that heavy alcohol drinking is commonly associated with an unhealthy lifestyle and meager nutrition (53), which might exert confounding effects that could not be ruled out in previous observational studies. Another possible reason is that the U-shaped association could not be detected in MR analysis. For example, light drinking may be associated with decreased risk of these diseases (11). In addition, it is also possible that the null associations observed in present study might be a consequence of inadequate power given SNPs used to mimic alcohol consumption explained a small phenotypic variance. In agreement with previous studies, our MR investigation demonstrated no associations of alcohol consumption with the development of gastroesophageal reflux, Crohn's disease, or ulcerative colitis (4, 6, 44).

Many mechanisms have been proposed to support the observed positive associations between smoking and gastrointestinal diseases. Tobacco smoking has

been shown to augment the production of numerous pro-inflammatory cytokines and decrease the levels of anti-inflammatory cytokines (54), which might mediate a variety of inflammation-associated gastrointestinal diseases. In addition, smoking may also generate impacts on the immune system, including inhibition of the function of circulatory dendritic cells (55) and alteration signaling of Toll-like receptors (56), which might contribute to the autoimmune disease and occurrence of neoplasm. The underlying mechanisms behind the associations of alcohol consumption with gastrointestinal diseases have not been fully understood. In addition to direct mucosal damage, the metabolites of ethanol are accountable for a part of the inflammation of alcohol drinking on the liver (57) and the gastrointestinal tract (58).

This study investigated the impacts of smoking and alcohol consumption on a wide range of gastrointestinal disease. Based on our findings, promoting public awareness of the adverse impacts of tobacco smoking and alcohol consumption on gastrointestinal diseases is of particular importance and should be used as prevention strategies to lower gastrointestinal disease burden because these two factors are modifiable behavioral factors as possible targets of the pharmacal (59) and behavioral interventions. In addition, our results may help facilitate the guidelines of gastrointestinal disease prevention and the management of certain patients who have a subsequent high risk of gastrointestinal disease., like those with obesity and diabetes (60, 61).

The major strength of the present study is MR design, which minimized bias from confounding and reverse causality and thus improved the causal inference in the associations of smoking and alcohol consumption with gastrointestinal diseases. We also used several independent outcome sources and combined the estimates, which increased statistical power as well as strengthened our findings by the observed consistency of results. Another strength is that we confined our analysis within the individuals of European ancestry, which minimized the population stratification bias.

This study also has several limitations. A major limitation of MR design is horizontal pleiotropy, which means that the used SNPs exerts effects on the outcomes not via the exposure but via alternative pathways. However, in this study, the bias caused by pleiotropic effects should be minimal since we observed no indications of horizontal pleiotropy in MR-Egger analysis, consistent results from a series of sensitivity analyses, and robust associations from multivariable MR analysis with mutual adjustment. Another limitation is the relatively small phenotypic variance of alcohol consumption (approximately 0.2%), which resulted in inadequate power to detect weak associations for certain uncommon gastrointestinal diseases. There are several limitations of using summary-level data. First, we could not evaluate the nonlinear associations between alcohol consumption and gastrointestinal diseases without individual-level data. We could not differentiate the associations of smoking and

alcohol consumption on the pathological subtypes of certain gastroenterological diseases, like esophageal cancer, based on summary-level data. For example, heavy alcohol consumption was associated with a high risk of squamous esophageal cancer (62), but the associations were inconsistent for adenocarcinoma esophageal cancer (63), which needs further investigation. Stratification analysis on sex was unlikely to be performed. In addition, we could not interpret and rescale the associations in a comparable scale to traditional observational studies because the unit of the exposure phenotypes were fixed in the corresponding genome-wide association analyses. An additional limitation is that our analysis was confined to the European populations, and thus whether the observed associations can be generalized to other populations remains unknown. For alcohol consumption, it has been reported that there were substantial behavioral and genetic differences across ethnic groups. For example, East Asian individuals drink much less alcohol compared to other races, which appears to be related to *ALDH2* gene (64). A further potential limitation is that the UK Biobank study were included in both the exposure and outcome datasets, which might cause MR estimates towards the observational associations. However, the used instrumental variants were proven to be strongly associated with the exposure (*F*-statistic > 10) (65), and the associations were replicated in the FinnGen study. Moreover, the associations remained stable in the sensitivity analyses using the genetic associations with exposures from the data

excluding the UK Biobank and 23andMe studies. All of these indicated that the bias due to sample overlap was limited.

In conclusion, this MR study suggested that smoking is a risk factor for a broad range of gastrointestinal diseases independent of alcohol consumption. Alcohol consumption on the other hand seemed to be an independent risk factor for only a few gastrointestinal diseases, including alcoholic liver disease, cirrhosis, and chronic pancreatitis, but we cannot rule out weak associations with other diseases. These findings provide genetic evidence on supporting reducing tobacco smoking and possibly excessive alcohol consumption in particular to prevent gastrointestinal diseases.

Declarations

Availability of data and materials

Data analyzed in the current study are publicly available GWAS summary-level data. The specific information and link could be found in Table S1. The code and curated data for the current analysis are available at https://github.com/XixianRuan/smoking_gi.

Funding

XL: the Natural Science Fund for Distinguished Young Scholars of Zhejiang Province (LR22H260001). XYW: National Natural Science Foundation of China (81970494) and Key Project of Research and Development Plan of Hunan Province(2019SK2041); SCL: the Swedish Heart Lung Foundation (Hjärt-Lungfonden, 20210351), the Swedish Research Council (Vetenskapsrådet, 2019-00977), and the Swedish Cancer Society (Cancerfonden). Funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Acknowledgement

We want to thank the Lee Lab, the FinnGen study, the International Inflammatory Bowel Disease Genetics Consortium (IIBDGC), and the Genetic Epidemiology Research on Aging (GERA) for sharing data.

Conflict of interest

All authors declare no competing interest.

References

1. Collaborators GBDA. Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2018;392(10152):1015-35.
2. Collaborators GBDRF. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1223-49.
3. Peery AF, Crockett SD, Murphy CC, Jensen ET, Kim HP, Egberg MD, et al. Burden and Cost of Gastrointestinal, Liver, and Pancreatic Diseases in the United States: Update 2021. Gastroenterology. 2022;162(2):621-44.
4. Eusebi LH, Ratnakumaran R, Yuan Y, Solaymani-Dodaran M, Bazzoli F, Ford AC. Global prevalence of, and risk factors for, gastro-oesophageal reflux symptoms: a meta-analysis. Gut. 2018;67(3):430-40.
5. Fund WCR, Research AIfC. Food, nutrition, physical activity, and the prevention of cancer: a global perspective: Amer Inst for Cancer Research; 2007.
6. Piovani D, Danese S, Peyrin-Biroulet L, Nikolopoulos GK, Lytras T, Bonovas S. Environmental Risk Factors for Inflammatory Bowel Diseases: An Umbrella Review of Meta-analyses. Gastroenterology. 2019;157(3):647-59 e4.
7. McGee EE, Jackson SS, Petrick JL, Van Dyke AL, Adami HO, Albanes D, et al. Smoking, Alcohol, and Biliary Tract Cancer Risk: A Pooling Project of 26 Prospective Studies. J Natl Cancer Inst. 2019;111(12):1263-78.
8. Yadav D, Whitcomb DC. The role of alcohol and smoking in pancreatitis. Nat Rev Gastroenterol Hepatol. 2010;7(3):131-45.
9. Bujanda L. The effects of alcohol consumption upon the gastrointestinal tract. Am J Gastroenterol. 2000;95(12):3374-82.
10. Laszkowska M, Rodriguez S, Kim J, Hur C. Heavy Alcohol Use Is Associated With Gastric Cancer: Analysis of the National Health and Nutrition Examination Survey From 1999 to 2010. Am J Gastroenterol. 2021;116(5):1083-6.
11. McNabb S, Harrison TA, Albanes D, Berndt SI, Brenner H, Caan BJ, et al. Meta-analysis of 16 studies of the association of alcohol with colorectal cancer. Int J Cancer. 2020;146(3):861-73.
12. Simpson RF, Hermon C, Liu B, Green J, Reeves GK, Beral V, et al. Alcohol drinking patterns and liver cirrhosis risk: analysis of the prospective UK Million Women Study. Lancet Public Health. 2019;4(1):e41-e8.
13. Roberts W, Verplaetse T, Peltier MKR, Moore KE, Gueorguieva R, McKee SA. Prospective association of e-cigarette and cigarette use with alcohol use in two waves of the Population Assessment of Tobacco and Health. Addiction. 2020;115(8):1571-9.
14. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet. 2019;51(2):237-44.
15. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89-98.
16. Yuan S, Larsson SC. Adiposity, diabetes, lifestyle factors and risk of gastroesophageal reflux

disease: a Mendelian randomization study. Eur J Epidemiol. 2022.
17. Larsson SC, Carter P, Kar S, Vithayathil M, Mason AM, Michaëlsson K, et al. Smoking, alcohol consumption, and cancer: A mendelian randomisation study in UK Biobank and international genetic consortia participants. PLoS Med. 2020;17(7):e1003178.
18. Yuan S, Larsson SC. Genetically Predicted Adiposity, Diabetes, and Lifestyle Factors in Relation to Diverticular Disease. Clin Gastroenterol Hepatol. 2022;20(5):1077-84.
19. Yuan S, Chen J, Li X, Fan R, Arsenault B, Gill D, et al. Lifestyle and metabolic factors for nonalcoholic fatty liver disease: Mendelian randomization study. Eur J Epidemiol. 2022.
20. Chen L, Yang H, Li H, He C, Yang L, Lv G. Insights into modifiable risk factors of cholelithiasis: A Mendelian randomization study. Hepatology. 2022;75(4):785-96.
21. Yuan S, Giovannucci EL, Larsson SC. Gallstone disease, diabetes, calcium, triglycerides, smoking and alcohol consumption and pancreatitis risk: Mendelian randomization study. NPJ Genom Med. 2021;6(1):27.
22. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779.
23. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner K, et al. FinnGen: Unique genetic insights from combining isolated population and national health register data. medRxiv. 2022:2022.03.03.22271360.
24. Wootton RE, Richmond RC, Stuijfzand BG, Lawn RB, Sallis HM, Taylor GMJ, et al. Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. Psychol Med. 2020;50(14):2435-43.
25. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015;47(9):979-86.
26. Guindo-Martinez M, Amela R, Bonas-Guarch S, Puiggros M, Salvoro C, Miguel-Escalada I, et al. The impact of non-additive genetic associations on age-related complex diseases. Nat Commun. 2021;12(1):2436.
27. Zhou W, Zhao Z, Nielsen JB, Fritsche LG, LeFaive J, Gagliano Taliun SA, et al. Scalable generalized linear mixed model for region-based association tests in large biobanks and cohorts. Nat Genet. 2020;52(6):634-9.
28. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017;46(6):1734-9.
29. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017;32(5):377-89.
30. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018;50(5):693-8.
31. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol. 2013;42(5):1497-501.
32. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018;7.

33. Castro C, Peleteiro B, Lunet N. Modifiable factors and esophageal cancer: a systematic review of published meta-analyses. J Gastroenterol. 2018;53(1):37-51.
34. Kato I, Nomura AM, Stemmermann GN, Chyou PH. A prospective study of gastric and duodenal ulcer and its relation to smoking, alcohol, and diet. Am J Epidemiol. 1992;135(5):521-30.
35. Nordenstedt H, Graham DY, Kramer JR, Rugge M, Verstovsek G, Fitzgerald S, et al. Helicobacter pylori-negative gastritis: prevalence and risk factors. Am J Gastroenterol. 2013;108(1):65-71.
36. Zhang YB, Pan XF, Chen J, Cao A, Zhang YG, Xia L, et al. Combined lifestyle factors, incident cancer, and cancer mortality: a systematic review and meta-analysis of prospective cohort studies. Br J Cancer. 2020;122(7):1085-93.
37. Talley NJ, Powell N, Walker MM, Jones MP, Ronkainen J, Forsberg A, et al. Role of smoking in functional dyspepsia and irritable bowel syndrome: three random population-based studies. Aliment Pharmacol Ther. 2021;54(1):32-42.
38. Aune D, Sen A, Leitzmann MF, Tonstad S, Norat T, Vatten LJ. Tobacco smoking and the risk of diverticular disease - a systematic review and meta-analysis of prospective studies. Colorectal Dis. 2017;19(7):621-33.
39. Liu B, Balkwill A, Roddam A, Brown A, Beral V, Million Women Study C. Separate and joint effects of alcohol and smoking on the risks of cirrhosis and gallbladder disease in middle-aged women. Am J Epidemiol. 2009;169(2):153-60.
40. Aune D, Vatten LJ, Boffetta P. Tobacco smoking and the risk of gallbladder disease. Eur J Epidemiol. 2016;31(7):643-53.
41. Aune D, Mahamat-Saleh Y, Norat T, Riboli E. Tobacco smoking and the risk of pancreatitis: A systematic review and meta-analysis of prospective studies. Pancreatology. 2019;19(8):1009-22.
42. Montgomery SM, Pounder RE, Wakefield AJ. Smoking in adults and passive smoking in children are associated with acute appendicitis. Lancet. 1999;353(9150):379.
43. Park S, Chun J, Han KD, Soh H, Kang EA, Lee HJ, et al. Dose-response relationship between cigarette smoking and risk of ulcerative colitis: a nationwide population-based study. J Gastroenterol. 2019;54(10):881-90.
44. Georgiou AN, Ntritsos G, Papadimitriou N, Dimou N, Evangelou E. Cigarette Smoking, Coffee Consumption, Alcohol Intake, and Risk of Crohn's Disease and Ulcerative Colitis: A Mendelian Randomization Study. Inflamm Bowel Dis. 2021;27(2):162-8.
45. Botteri E, Borroni E, Sloan EK, Bagnardi V, Bosetti C, Peveri G, et al. Smoking and Colorectal Cancer Risk, Overall and by Molecular Subtypes: A Meta-Analysis. Am J Gastroenterol. 2020;115(12):1940-9.
46. Dimou N, Yarmolinsky J, Bouras E, Tsilidis KK, Martin RM, Lewis SJ, et al. Causal Effects of Lifetime Smoking on Breast and Colorectal Cancer Risk: Mendelian Randomization Study. Cancer Epidemiol Biomarkers Prev. 2021;30(5):953-64.
47. Mizrahi JD, Surana R, Valle JW, Shroff RT. Pancreatic cancer. Lancet. 2020;395(10242):2008-20.
48. Yu C, Tang H, Guo Y, Bian Z, Yang L, Chen Y, et al. Hot Tea Consumption and Its Interactions With Alcohol and Tobacco Use on the Risk for Esophageal Cancer: A Population-Based Cohort Study. Ann Intern Med. 2018;168(7):489-97.

49. Roerecke M, Vafaei A, Hasan OSM, Chrystoja BR, Cruz M, Lee R, et al. Alcohol Consumption and Risk of Liver Cirrhosis: A Systematic Review and Meta-Analysis. Am J Gastroenterol. 2019;114(10):1574-86.
50. Samokhvalov AV, Rehm J, Roerecke M. Alcohol Consumption as a Risk Factor for Acute and Chronic Pancreatitis: A Systematic Review and a Series of Meta-analyses. EBioMedicine. 2015;2(12):1996-2002.
51. Ryan-Harshman M, Aldoori W. How diet and lifestyle affect duodenal ulcers. Review of the evidence. Can Fam Physician. 2004;50:727-32.
52. Bagnardi V, Rota M, Botteri E, Tramacere I, Islami F, Fedirko V, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. Br J Cancer. 2015;112(3):580-93.
53. Klatsky AL. Diet, alcohol, and health: a story of connections, confounders, and cofactors. Am J Clin Nutr. 2001;74(3):279-80.
54. Arnson Y, Shoenfeld Y, Amital H. Effects of tobacco smoke on immunity, inflammation and autoimmunity. J Autoimmun. 2010;34(3):J258-65.
55. Givi ME, Folkerts G, Wagenaar GT, Redegeld FA, Mortaz E. Cigarette smoke differentially modulates dendritic cell maturation and function in time. Respir Res. 2015;16:131.
56. Noakes PS, Hale J, Thomas R, Lane C, Devadason SG, Prescott SL. Maternal smoking is associated with impaired neonatal toll-like-receptor-mediated immune responses. Eur Respir J. 2006;28(4):721-9.
57. Mandrekar P, Szabo G. Signalling pathways in alcohol-induced liver inflammation. J Hepatol. 2009;50(6):1258-66.
58. Bishehsari F, Magno E, Swanson G, Desai V, Voigt RM, Forsyth CB, et al. Alcohol and Gut-Derived Inflammation. Alcohol Res. 2017;38(2):163-71.
59. Leone FT, Zhang Y, Evers-Casey S, Evins AE, Eakin MN, Fathi J, et al. Initiating Pharmacologic Treatment in Tobacco-Dependent Adults. An Official American Thoracic Society Clinical Practice Guideline. Am J Respir Crit Care Med. 2020;202(2):e5-e31.
60. Camilleri M, Malhi H, Acosta A. Gastrointestinal Complications of Obesity. Gastroenterology. 2017;152(7):1656-70.
61. Krishnan B, Babu S, Walker J, Walker AB, Pappachan JM. Gastrointestinal complications of diabetes mellitus. World J Diabetes. 2013;4(3):51-63.
62. Abnet CC, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma. Gastroenterology. 2018;154(2):360-73.
63. Coleman HG, Xie SH, Lagergren J. The Epidemiology of Esophageal Adenocarcinoma. Gastroenterology. 2018;154(2):390-405.
64. Jorgenson E, Thai KK, Hoffmann TJ, Sakoda LC, Kvale MN, Banda Y, et al. Genetic contributors to variation in alcohol consumption vary by race/ethnicity in a large multi-ethnic genome-wide association study. Mol Psychiatry. 2017;22(9):1359-67.
65. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol. 2016;40(7):597-608.

Table and figure legends

Table 1. Associations of genetic predisposition to smoking initiation with 24 gastrointestinal diseases in univariable and multivariable Mendelian randomization analyses.

Table 2. Associations of genetically predicted alcohol consumption with 24 gastrointestinal diseases in univariable and multivariable Mendelian randomization analyses.

Fig 1. Overview of the present study design. GERA, Genetic Epidemiology Research on Aging; IIBDGC, the International Inflammatory Bowel Disease Genetics Consortium; MR, Mendelian randomization; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier; SNP, single nucleotide polymorphism.

Fig 2. Summary of associations of genetically predicted smoking initiation, lifetime smoking, and alcohol consumption with 24 gastrointestinal diseases. UVMR, univariable Mendelian randomization; MVMR, multivariable Mendelian randomization. The numbers in the box are the odds ratios for associations of exposure for gastrointestinal diseases. The association with a *P*-value <0.05 but Benjamini-Hochberg adjusted *P*-value >0.05 was regarded suggestive and the association with a Benjamini–Hochberg adjusted P-value <0.05 were deemed significant.

Table 1. Associations of genetic predisposition to smoking initiation with 24 gastrointestinal diseases in univariable and multivariable Mendelian randomization analyses.

Disease		Total cases	Total controls	UVMR			MVMR adjusted for alcohol consumption	
				OR (95% CI)	P value	I^2 (95% CI)	OR (95% CI)	P value
Upper Gastrointestinal Diseases	Gastroesophageal reflux	34,135	634,629	1.28 (1.20, 1.37)	4.09×10^{-14}*	46.24	1.65 (1.35, 2.02)	1.38×10^{-6}*
	Esophageal cancer	1,130	702,116	1.67 (1.24, 2.25)	6.84×10^{-4}*	22.68	4.78 (2.10, 10.90)	1.97×10^{-4}*
	Gastric ulcer	8,651	666,879	1.54 (1.37, 1.72)	3.83×10^{-14}*	44.96	1.95 (1.40, 2.71)	7.31×10^{-5}*
	Duodenal ulcer	5,713	666,879	1.53 (1.34, 1.75)	8.47×10^{-10}*	0.00	1.64 (1.07, 2.52)	0.024
	Acute Gastritis	3,048	643,478	1.29 (1.09, 1.53)	0.003*	0.00	1.54 (0.91, 2.62)	0.106
	Chronic gastritis	7,975	643,478	1.33 (1.18, 1.49)	1.55×10^{-6}*	77.04	1.33 (0.96, 1.86)	0.091
	Gastric cancer	1,608	701,472	1.42 (1.13, 1.79)	0.002*	0.00	2.29 (1.14, 4.59)	0.020
Lower Gastrointestinal Diseases	Irritable bowel disease	15,718	641,489	1.22 (1.12, 1.32)	3.50×10^{-6}*	11.84	1.43 (1.10, 1.85)	0.008*
	Celiac disease	4,808	631,700	0.82 (0.66, 1.02)	0.071	0.00	0.87 (0.53, 1.43)	0.59
	Diverticular disease	50,065	587,969	1.25 (1.18, 1.33)	5.23×10^{-14}*	67.29	1.56 (1.30, 1.87)	1.41×10^{-6}*
	Crohn's disease	10,846	645,718	1.25 (1.11, 1.40)	3.03×10^{-4}*	0.00	1.48 (1.01, 2.16)	0.042
	Ulcerative colitis	16,770	651,255	1.15 (1.04, 1.26)	0.004*	0.00	0.94 (0.71, 1.25)	0.677
	Colorectal cancer	9,519	686,953	1.03 (0.92, 1.14)	0.632	29.94	1.03 (0.76, 1.39)	0.841
Hepatobiliary and Pancreatic Diseases	Non-alcoholic fatty liver disease	3,242	707,631	1.49 (1.26, 1.76)	3.82×10^{-6}*	0.00	2.11 (1.15, 3.88)	0.016*
	Alcoholic liver disease	2,955	680,369	1.99 (1.65, 2.41)	1.49×10^{-12}*	92.68	2.26 (1.26, 4.03)	0.006
	Cirrhosis	5,904	706,200	1.68 (1.40, 2.02)	3.39×10^{-8}*	0.00	1.92 (1.06, 3.47)	0.032
	Liver cancer	714	702,008	1.57 (1.13, 2.17)	0.007*	0.00	1.96 (0.73, 5.25)	0.183
	Cholangitis	1,708	664,749	1.02 (0.80, 1.29)	0.892	0.00	1.31 (0.61, 2.84)	0.489
	Cholecystitis	5,893	664,749	1.47 (1.29, 1.68)	4.71×10^{-9}*	84.72	2.38 (1.57, 3.60)	4.14×10^{-5}*
	Cholelithiasis	42,510	664,749	1.20 (1.13, 1.27)	5.75×10^{-9}*	0.00	1.33 (1.02, 1.73)	0.035
	Acute pancreatitis	6,634	679,713	1.39 (1.23, 1.56)	6.71×10^{-8}*	79.71	1.55 (1.04, 2.31)	0.031
	Chronic pancreatitis	3,173	679,713	1.38 (1.17, 1.64)	1.79×10^{-4}*	0.00	1.27 (0.74, 2.16)	0.384
	Pancreatic cancer	1,643	701,472	1.00 (0.79, 1.26)	0.999	67.21	2.08 (1.06, 4.10)	0.034
Other	Acute appendicitis	25,361	690,149	1.15 (1.08, 1.23)	1.27×10^{-5}*	0.00	1.15 (0.92, 1.44)	0.221

UVMR, univariable Mendelian randomization; MVMR, multivariable Mendelian randomization; OR, odds ratio; CI, confidence interval. *Significant association after multiple testing.

Table 2. Associations of genetically predicted alcohol consumption with 24 gastrointestinal diseases in univariable and multivariable Mendelian randomization analyses.

Disease		Total cases	Total controls	UVMR			MVMR adjusted for smoking initiation	
				OR (95% CI)	P value	I^2 (95% CI)	OR (95% CI)	P value
Upper Gastrointestinal Diseases	Gastroesophageal reflux	34,135	634,629	0.99 (0.81, 1.21)	0.893	46.24	0.88 (0.72, 1.08)	0.219
	Esophageal cancer	1,130	702,116	2.86 (1.18, 6.91)	0.020	22.68	1.28 (0.59, 2.82)	0.533
	Gastric ulcer	8,651	666,879	1.30 (0.95, 1.77)	0.098	44.96	1.06 (0.77, 1.47)	0.721
	Duodenal ulcer	5,713	666,879	1.92 (1.23, 3.00)	0.004*	0.00	1.54 (1.01, 2.34)	0.045
	Acute Gastritis	3,048	643,478	0.99 (0.58, 1.69)	0.960	0.00	0.88 (0.52, 1.48)	0.621
	Chronic gastritis	7,975	643,478	1.33 (0.90, 1.95)	0.147	77.04	1.33 (0.93, 1.89)	0.115
	Gastric cancer	1,608	701,472	1.57 (0.75, 3.30)	0.233	0.00	1.59 (0.79, 3.21)	0.194
Lower Gastrointestinal Diseases	Irritable bowel disease	15,718	641,489	0.73 (0.57, 0.93)	0.012	11.84	0.74 (0.57, 0.97)	0.027
	Celiac disease	4,808	631,700	0.87 (0.53, 1.42)	0.097	0.00	1.04 (0.64, 1.68)	0.887
	Diverticular disease	50,065	587,969	0.95 (0.79, 1.13)	0.553	67.29	0.94 (0.79, 1.13)	0.527
	Crohn's disease	10,846	645,718	0.91 (0.62, 1.32)	0.613	0.00	0.74 (0.53, 1.05)	0.088
	Ulcerative colitis	16,770	651,255	1.11 (0.82, 1.50)	0.509	0.00	0.88 (0.67, 1.15)	0.358
	Colorectal cancer	9,519	686,953	1.09 (0.76, 1.55)	0.649	29.94	1.28 (0.95, 1.72)	0.098
Hepatobiliary and Pancreatic Diseases	Non-alcoholic fatty liver disease	3,242	707,631	1.20 (0.63, 2.28)	0.574	0.00	0.99 (0.54, 1.79)	0.962
	Alcoholic liver disease	2,955	680,369	14.35 (7.69, 26.81)	6.32×10^{-17}*	92.68	9.60 (5.28, 17.46)	1.25×10^{-13}*
	Cirrhosis	5,904	706,200	2.96 (1.50, 5.85)	0.002*	0.00	2.41 (1.29, 4.52)	0.006*
	Liver cancer	714	702,008	1.16 (0.43, 3.11)	0.775	0.00	0.76 (0.29, 2.02)	0.585
	Cholangitis	1,708	664,749	0.96 (0.44, 2.08)	0.912	0.00	0.72 (0.33, 1.55)	0.397
	Cholecystitis	5,893	664,749	1.36 (0.91, 2.03)	0.132	84.72	0.96 (0.64, 1.45)	0.862
	Cholelithiasis	42,510	664,749	1.02 (0.75, 1.39)	0.878	0.00	1.03 (0.79, 1.35)	0.801
	Acute pancreatitis	6,634	679,713	1.36 (0.91, 2.03)	0.128	79.71	1.17 (0.78, 1.75)	0.456
	Chronic pancreatitis	3,173	679,713	2.96 (1.80, 4.89)	2.13×10^{-5}*	0.00	3.24 (1.86, 5.64)	3.18×10^{-5}*
	Pancreatic cancer	1,643	701,472	0.63 (0.32, 1.26)	0.193	67.21	0.79 (0.40, 1.56)	0.496
Other	Acute appendicitis	25,361	690,149	0.80 (0.63, 1.01)	0.063	0.00	0.77 (0.61, 0.97)	0.024

UVMR, univariable Mendelian randomization; MVMR, multivariable Mendelian randomization; OR, odds ratio; CI, confidence interval.

*Significant association after multiple testing