Food & Function

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Food Funct.*, 2024, **15**, 1705

Causal association between tea consumption and head and neck cancer: a Mendelian randomization study†

Qi-he Zhang,[a] Mei-qi Wang,[b] Huan-huan Wang,[a,c,d] Yu-wei Huang,[e] Chao Dong,[e] Ying Xin *‡[e] and Xin Jiang *‡[a,c,d]

Received 20th September 2023,
Accepted 4th January 2024

DOI: 10.1039/d3fo04017h

rsc.li/food-function

Although evidence supports an observational association between tea consumption and susceptibility to head and neck cancer, the causal nature of this association remains unclear. We performed a two-sample Mendelian randomization (MR) analysis to determine the causal effects of tea consumption on head and neck cancer. We employed a fixed-effects inverse variance-weighted model for the MR analysis. Genome-wide association study (GWAS) summary data for tea consumption were obtained from the UK Biobank Consortium, and GWAS data for head and neck cancer were derived from two data sources and were used as the outcomes. Our MR analysis revealed limited evidence for a causal relationship between various types of tea intake and head and neck cancer. After adjustment for smoking and alcohol consumption, there was no causal relationship between tea consumption and head and neck cancer. Further experimental studies are required to confirm its potential role in these malignancies.

Introduction

Head and neck cancer ranks as the sixth most common cancer worldwide,[1] with oral and oropharyngeal cancers being the most prevalent subtypes. Established risk factors for these cancers encompass tobacco and alcohol consumption,[2] human papillomavirus (HPV) infection,[3] and oral sexual behaviors.[4]

Tea, one of the world's most widely consumed beverages, encompasses various types such as green, black, herbal, and white tea. It contains a plethora of chemical compounds, including catechins, tea polyphenols, caffeine, theanine, amino acids, volatile oils, and minerals. Notably, catechins and tea polyphenols, the most abundant compounds in tea, exhibit diverse biological activities including antioxidant, anti-inflammatory, and anticancer properties. A meta-analysis has indicated an inverse association between tea consumption and the risk of 11 cancers including biliary tract, breast, colorectal, endometrial, and gastric cancers.[5] Nevertheless, some studies have yielded inconclusive results regarding tea's impact on cancer.[6] Given that tea consumption directly exposes the oral cavity, oropharynx, and larynx to its compounds, its potential role in protecting against tumors in these regions prompted our investigation into whether tea consumption has a protective effect against head and neck cancer. Mendelian randomization (MR) is a statistical method used to evaluate the causal relationship between exposure and outcome using instrumental variables (genetic variants). This approach can be considered a natural analog of randomized controlled trials that is devoid of confounding and reverse causality bias. In contrast to the traditional gold-standard randomized controlled trials for causal inference, patients are assigned according to their genotype, avoiding reverse causality bias and influence of confounding factors such as ethical and socioeconomic factors. Accordingly, we aimed to determine whether tea consumption is causally related to head and neck cancer using a two-sample MR analysis. We hypothesized that tea consumption may increase or decrease susceptibility to head and neck cancer. Our study aims to fill the current knowledge gap regarding the protective effects of tea against head and neck cancer and further validate previous research findings to provide a more comprehensive understanding.

[a]*Jilin Provincial Key Laboratory of Radiation Oncology & Therapy, The First Hospital of Jilin University, and Key Laboratory of Pathobiology, Ministry of Education, Jilin University, Changchun 130021, China*
[b]*Department of Gastrointestinal Colorectal and Anal Surgery, China-Japan Union Hospital of Jilin University, Changchun 130021, China*
[c]*Department of Radiation Oncology, The First Hospital of Jilin University, Changchun 130021, China*
[d]*NHC Key Laboratory of Radiobiology, School of Public Health, Jilin University, Changchun 130021, China*
[e]*Key Laboratory of Pathobiology, Ministry of Education, College of Basic Medical Sciences, Jilin University, Changchun 130021, China*

†Electronic supplementary information (ESI) available: Supplementary Fig. 1–26 and Tables 1–4. See DOI: https://doi.org/10.1039/d3fo04017h

‡These authors have contributed equally to this work and share corresponding authorship.

Methods

Study design and data sources

The genome-wide association study (GWAS) summarized the tea intake and head and neck cancer data, which were obtained from the MRC IEU OpenGWAS data infrastructure.[7] The original head and neck cancer data were obtained from the OncoArray Oral Cavity and Oropharyngeal Cancer Consortium, and FinnGen biobank[8,9] (Fig. 1). Outcome details are presented in Table 1. The original investigation was conducted after receiving ethical approval for each study included in the MR analysis. The exposure data were released from the UK Biobank. The UK Biobank began in 2006 by recruiting approximately 500 000 participants between the ages of 38 and 73. Participants completed a series of questionnaires that provided detailed personal and lifestyle information. In addition, participants provided biospecimens, including blood, urine, and saliva, which were subsequently sequenced for genome sequencing and genotyping using the Illumina sequencing platform in the UK Biobank. Details of tea consumption (UKB Data-Field: 1488) of the subjects, both male and female, were asked: "How many cups of tea do you drink each day? (Include black and green tea)". The sample size of tea intake was 447485. Units of measurement are cups per day. Mean = 3.494, Std dev. = 2.84157. The rest of the exposure information summary is given in ESI Table 1.†

Selection of instrumental variables

We extracted eight GWAS summary exposure datasets (Table 1). Single-nucleotide polymorphisms (SNPs) with a significance level within the locus-wide range (5×10^{-8}) were selected as instrumental variables, as used in a previous study. If effective instrumental variables could not be extracted, the p value was increased to 5×10^{-6}. One MR principle is that there should be no linkage disequilibrium (LD) between the included instrumental variables, as a strong LD may result in biased outcomes. In this study, clumping processing ($R^2 < 0.001$, clumping distance = 10 000 kb) was performed to evaluate the LD among the SNPs.[10]

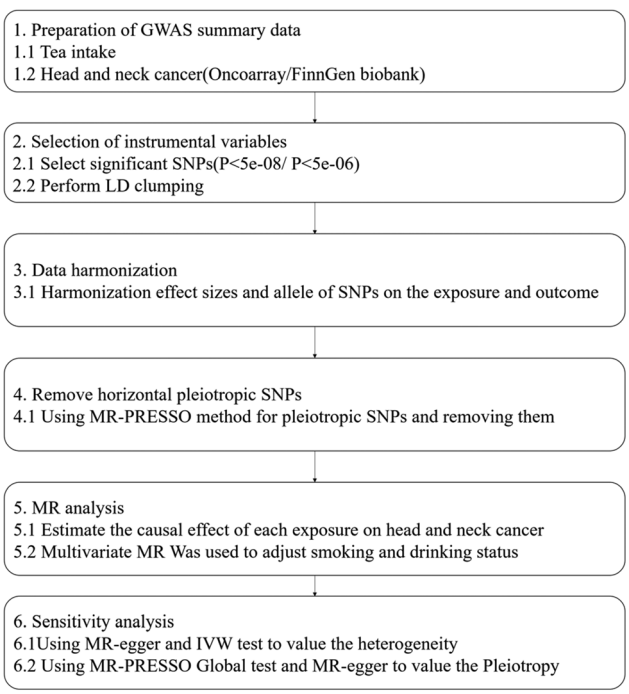

Fig. 1 The whole workflow of MR analysis. GWAS, genome-wide association study; SNPs, single nucleotide polymorphisms; LD, linkage disequilibrium; IVW, inverse variance weighted.

Assumptions

This two-sample MR study relied on three critical assumptions to minimize bias. First, the genetic instruments used were significantly associated with exposure. Second, the instrumental variables were independent of confounders that influenced both the exposure and outcomes. Third, instrumental variables affected the outcomes solely through exposure, implying no horizontal pleiotropy effect between the instrumental variables and outcomes.

Statistical methods

Various analytical methods, including inverse variance-weighted (IVW),[11] maximum likelihood,[12] MR-Egger,[13] weighted median,[14] weighted mode, and MR-PRESSO,[15] were employed to infer potential causality. The IVW results were considered robust in the absence of horizontal pleiotropy. The maximum likelihood method resembled the IVW method; however, it considered the uncertainty of the SNP-exposure association and the overlap of samples in the two-sample MR studies.[16] In MR-Egger's assumption, the presence of an intercept term was considered and used to assess pleiotropy. If this intercept term was close to zero, the MR-Egger regression model became similar to the IVW model. However, if the intercept term was different from zero, horizontal pleiotropy may have occurred among these instrumental variables (IVs). The weighted median can provide consistent estimates of causal effects, even if <50% of the SNPs have pleiotropy. When most instrumental variables did not meet the causal inference requirements of the MR method, the weighted model was considered valid. The IVW method has been reported to be slightly more robust than the other methods under certain conditions. Therefore, the results were primarily based on the IVW method and were supplemented by the other methods.

Furthermore, we conducted multivariate Mendelian randomization (MVMR) to account for potential confounding factors, employing Bonferroni's adjustment ($P = 0.05/3$). We performed an MVMR analysis using the mr_mvivw, mr_mvegger, and mr_mvlasso functions in the R package "MendelianRandomizaton". To mitigate potential issues of collinearity, we also carried out lasso regression as an additional supplementary analysis.

Colocalization analysis

Colocalization analysis is frequently employed to ascertain whether two phenotypes are influenced by an identical causal variant within a designated genomic region, thereby providing substantiating proof of a connection between these two pheno-

Table 1 Outcomes and exposure of GWAS samples used in this study

GWAS ID	Trait	Consortium	Sample size	Number of SNPs	Population	*N* case	*N* control	PMID
ieu-b-89	Oral cavity and pharyngeal cancer	Oncoarray oral cavity and oropharyngeal cancer	5425	7 514 278	European (Geographic region: Europe)	2497	2928	27749845
ieu-b-94	Oral cavity cancer	Oncoarray oral cavity and oropharyngeal cancer	4151	7 510 833	European (Geographic region: Europe)	1223	2928	27749845
ieu-b-96	Oropharyngeal cancer	Oncoarray oral cavity and oropharyngeal cancer	4018	7 508 444	European (Geographic region: Europe)	1090	2928	27749845
ieu-b-90	Oral cavity and pharyngeal cancer	Oncoarray oral cavity and oropharyngeal cancer	4671	7 510 261	European (Geographic region: North America)	2342	2329	27749845
ieu-b-93	Oral cavity cancer	Oncoarray oral cavity and oropharyngeal cancer	3464	7 506 142	European (Geographic region: North America)	1135	2329	27749845
ieu-b-97	Oropharyngeal cancer	Oncoarray oral cavity and oropharyngeal cancer	3448	7 506 485	European (Geographic region: North America)	1119	2329	
finn-b-C3_LIP_ORAL_PHARYNX	Malignant neoplasm of the lips, oral cavity and pharynx	FinnGen biobank	218 792	16 380 466	European	126	218 666	
finn-b-C3_LIP_ORAL_PHARYNX_EXALLC	Malignant neoplasm of the lips, oral cavity and pharynx (all cancers excluded)	FinnGen biobank	174 132	16 380 304	European	126	174 006	
finn-b-C3_LARYNX	Malignant neoplasm of the larynx	FinnGen biobank	218 792	16 380 466	European	180	218 612	
finn-b-C3_LARYNX_EXALLC	Malignant neoplasm of the larynx (all cancers excluded)	FinnGen biobank	174 185	16 380 304	European	180	174 005	
ukb-b-6066	Tea intake	MRC-IEU	447 485	9 851 867	European	Continuous		
ukb-b-4078	Green tea intake	MRC-IEU	64 949	9 851 867	European	Categorically ordered		
ukb-b-13344	Herbal tea intake	MRC-IEU	64 949	9 851 867	European	Categorically ordered		
ukb-b-17988	Tea consumed	MRC-IEU	64 949	9 851 867	European	Binary		
ukb-b-8553	Decaffeinated tea	MRC-IEU	64 949	9 851 867	European	Categorically ordered		
ukb-b-3291	Standard tea intake	MRC-IEU	64 949	9 851 867	European	Categorically ordered		
ukb-b-11491	Added milk to rooibos tea	MRC-IEU	64 949	9 851 867	European	Categorically ordered		
ukb-b-5209	Added milk to standard tea	MRC-IEU	64 949	9 851 867	European	Categorically ordered		

types. We gathered the tea characteristics and their corresponding SNPs that satisfied the genome-wide association study threshold (for ukb-b-8553, SNPs could not be obtained under this threshold, so it was adjusted to 5×10^{-7}) in preparation for the subsequent colocalization analysis. The "ideal" MR analysis we envision involves extracting instrumental variables after meeting the threshold for a genome-wide association study. Colocalization analysis serves as a supplementary analysis when an insufficient number of effective SNPs were obtained in genome-wide association studies. Due to ukb-b-3291, ukb-b-5209, ukb-b-8553, ukb-b-11491, and ukb-b-17988 not yielding a sufficient number of SNPs under the standard genome-wide association study threshold (5×10^{-8}), only the Wald ratio method could be used for MR. Therefore, we believe that supplementary colocalization analysis can better reflect the presence of causal effects. As for ukb-b-6066, ukb-b-4078, and ukb-b-13344, MR using IVW and similar methods can be performed under the genome-wide association study threshold (5×10^{-8}). We considered their ability to assess causal effects strong enough, eliminating the need for colocalization analysis as a supplement. After careful consideration, we did not conduct colocalization analysis for these three exposures.

Specifically, we employed SNP positions that meet the threshold criteria and fall within a 500 kb (on the analysis of the added milk to rooibos tea (id: ukb-b-11497) for the FinnGen biobank source outcome, unable to extract the corresponding SNPS, in the end will be expanded to 1000 kb) window both upstream and downstream as potential SNPs for extraction in terms of both exposure and outcome traits for the purpose of conducting a colocalization analysis. This analysis encompasses five distinct model assumptions, as follows: H0, where no significant association is present between all SNP loci within a genomic region and both the exposure and outcome; H1/H2, indicating a significant association between either the exposure or outcome and SNP loci within a genomic region; H3, denoting a significant association between both the exposure and outcome and SNP loci within a genomic region, driven by distinct causal variants; and H4, representing a significant association between both the exposure and outcome and SNP loci within a genomic region, driven by the same causal variant. During the course of colocalization analysis, posterior probabilities (PP.H0–PP.H4) were computed for each of these models, and the sum of these posterior probabilities for the five models equals 1. A higher posterior probability associated with a specific model indicates a greater likelihood of that model assumption being valid based on the data. We consider the H4 model assumption as valid when PP.H4 > 0.80.[17]

Assessment of assumptions

We estimated the variance of each tea intake. The power of our MR analyses was assessed using the online calculator mRnd (**https://shiny.cnsgenomic.com**).

Sensitivity analyses

We tested for heterogeneity using the mr_heterogeneity function in the R package "TwoSampleMR"; the mr_heterogeneity function was performed using Cochran's Q test in the IVW test and MR-Egger regression. Horizontal pleiotropy was tested using the mr_pleiotropy_test function in the R package "TwoSampleMR", which uses the MR Egger method. MR-Egger regression was used to estimate the effect of pleiotropy, yielding a more robust pleiotropy-corrected causal estimate under the assumptions of no measurement error and instrument strength independent of direct effects.[18] If MR-Egger detected the presence of pleiotropy, MR-PRESSO[19] was used to correct outliers. Leave-one-out analysis was used to ascertain whether a single SNP exerted causal effects. We also assessed the instrument strength using the F statistic,[20] calculated using the following formula:

$$F = \frac{R^2(N-k-1)}{k(1-R^2)} \quad (1)$$

where R^2 represents the variance in exposure explained by the selected SNPs, N is the sample size, and k represents the number of instrumental variables. If $F < 10$, indicating a higher likelihood of weak instrumental bias, the association between the instrumental variables and exposure was considered weak.

R square for each SNP was calculated using the following formula:[21,22]

$$R^2 = \frac{(2 \times \mathrm{EAF} \times (1-\mathrm{EAF}) \times \mathrm{beta}^2)}{(2 \times \mathrm{EAF} \times (1-\mathrm{EAF}) \times \mathrm{beta}^2) + (2 \times \mathrm{EAF} \times (1-\mathrm{EAF}) \times N \times SE^2)} \quad (2)$$

In this context, EAF represents the frequency of the effect allele, beta signifies the estimated genetic impact on exposure, N stands for the sample size of the GWAS concerning the association between the SNP and exposure, and SE represents the standard error of the genetic effect.

We used the mr_mvivw and pleiotropy_mvmr functions in the R package "Mendelian Randomization" to evaluate heterogeneity and pleiotropy in MVMR. In addition, our IVW method used a random-effects model to exclude the interference of heterogeneity on the results.

Software and pre-registration

Analysis was performed using R software (version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria) and R packages "TwoSampleMR" (version 0.5.6),[23] and "MRPRESSO".[15] We adhered to the STROBE-MR guidelines for strengthening the reporting of observational studies in epidemiological studies using MR to report our results.[24]

Results

Selection of instrumental variables

After removing palindromic SNPs, performing clumping, and harmonizing the data, the SNPs associated with all outcomes ranged from 1 to 74. The R^2 and F statistic values for the exposures are summarized in Table S1.† The F statistic values for all exposures were >10, indicating no evidence of weak

instrumental bias and demonstrating that all SNPs had sufficient validity.

Univariate MR analysis

We conducted a primary MR analysis (under genome-wide association study threshold criteria 5×10^{-8}) to investigate the causal effects of 8 tea intake-related exposures on 10 outcomes. The results indicated that green tea intake had a causal effect on the development of oral cancer in the OncoArray (European population) (OR = 0.936, [range: 0.878–0.999], p = 0.046) (Fig. 2–5 and Table S1†).

Subsequently, we conducted a colocalization analysis of the exposure (standard tea intake || id: ukb-b-3291; added milk to standard tea || id: ukb-b-5209; decaffeinated tea || id: ukb-b-8553; added milk to rooibos tea || id: ukb-b-11491; tea consumed || id: ukb-b-17988) and the outcome, for which a valid number of SNPs could not be extracted under the threshold criteria for the initial genome-wide association study. The SNPs used for colocalization analysis, as well as the number of SNPs within the 500 kb upstream and downstream regions, can be found in ESI Table S2.† These were extracted according to the threshold criteria for genome-wide association studies (5×10^{-8}), except for decaffeinated tea (GWAS id: ukb-b-8553); these SNPs could not be obtained at this threshold, so they were adjusted to 5×10^{-7}, as described in the Methods section.

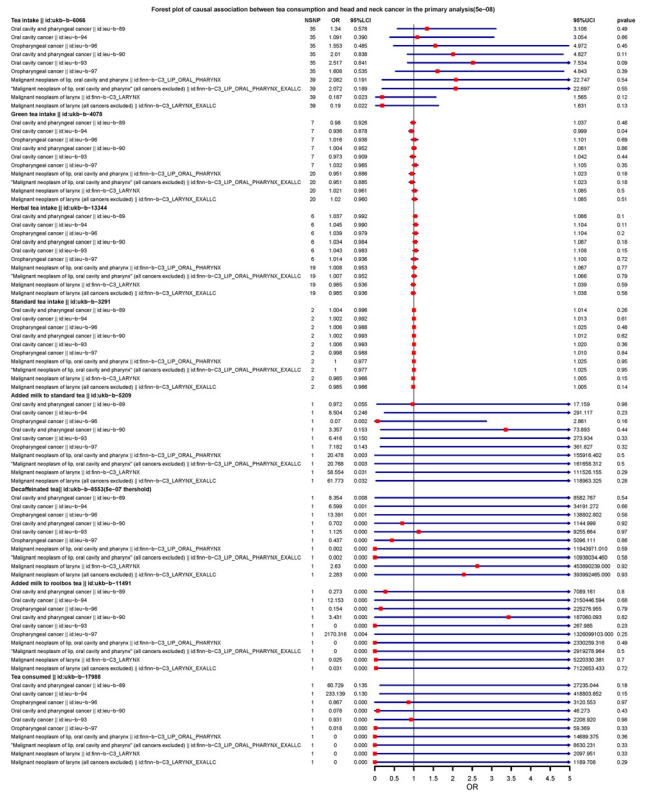

Fig. 2 Forest plot of causal association between tea consumption and head and neck cancer in the primary analysis (5×10^{-8}). NSNP, number of SNPs; OR, odds ratio; IVW, inverse variance weighted; CI, confidence interval.

The results showed that none of the PP.H4 values exceeded 0.80, providing no evidence to support the hypothesis that the exposure and outcome are significantly associated with SNP sites within genomic regions driven by the same pathogenic variant (Table S2†).

In our secondary MR analysis, we have adjusted the threshold for instrumental variables to a more lenient criterion (5×10^{-6}). The results indicate that added milk to rooibos tea had causal effects on the occurrence of malignant neoplasm of the lips, oral cavity and pharynx (with all cancers excluded or not) (OR = 9.2968×10^{-6}, [range: 2.87597×10^{-9}–0.030], p = 0.005), and malignant neoplasm of the lips, oral cavity and pharynx (with all cancers excluded) (OR = 8.86985×10^{-6}, [range: 2.7518×10^{-9}–0.029], p = 0.005). None of the remaining exposures demonstrated causal effects on head and neck cancer (Table S1†) (Fig. 6).

Multivariate MR analysis

In order to account for potential confounding factors, we conducted multivariable MR analyses under different thresholds for instrumental variable extraction (5×10^{-8}, 5×10^{-7}). For three exposures (tea intake || id: ukb-b-6066; green tea intake || id: ukb-b-4078; herbal tea intake || id: ukb-b-13344) for which a sufficient number of SNPs could be obtained in genome-wide association studies, we performed multivariable MR using instrumental variables extracted under the 5×10^{-8} threshold. For the remaining five exposures for which instrumental variables could not be extracted under the 5×10^{-8} threshold, we initially used the 5×10^{-7} threshold for instrumental variable extraction, with the aim of providing a more comprehensive perspective by lowering the threshold criteria. Smoking and drinking are high-risk factors of head and neck tumors, and tea intake is associated with smoking and drinking. Therefore, we searched for all phenotypes related to smoking and drinking in the IEU open GWAS for the univariate MR analysis. The results revealed that smoking initiation and alcoholic drinks per week had causal effects on the outcome of head and neck tumors from seven and nine different data sources, respectively; consequently, these two exposures were included in the multivariate MR analysis.

The multivariate MR analysis indicated that green tea intake had a protective effect on oral cancer (from the North American population of Oncoarray) (OR_{IVW} = 0.955, [range: 0.910–0.999], p_{IVW} = 0.041) after lasso regression; the results were still statistically significant. As for tea intake (id: ukb-b-6066), it has a causal effect on malignant neoplasm of the larynx and malignant neoplasm of the larynx (all cancers excluded) ($OR_{MR\text{-}Egger}$ = 0.064, [range: 0.004–0.980], $p_{MR\text{-}Egger}$ = 0.048), ($OR_{MR\text{-}Egger}$ = 0.060, [range: 0.004–0.925], $p_{MR\text{-}Egger}$ = 0.044). Adding milk to standard tea was a risk factor (FinnGen Biobank) against both malignant neoplasm of the larynx and malignant neoplasm of the larynx (all cancers excluded) (OR_{lasso} = 11.533, [range: 1.449–91.768], p = 0.021) (OR_{lasso} = 12.254, [range: 1.544–97.273], p = 0.018). However, the causal effect was pleiotropic, and the MR-Egger results showed that the causal effect was not statistically significant. For added

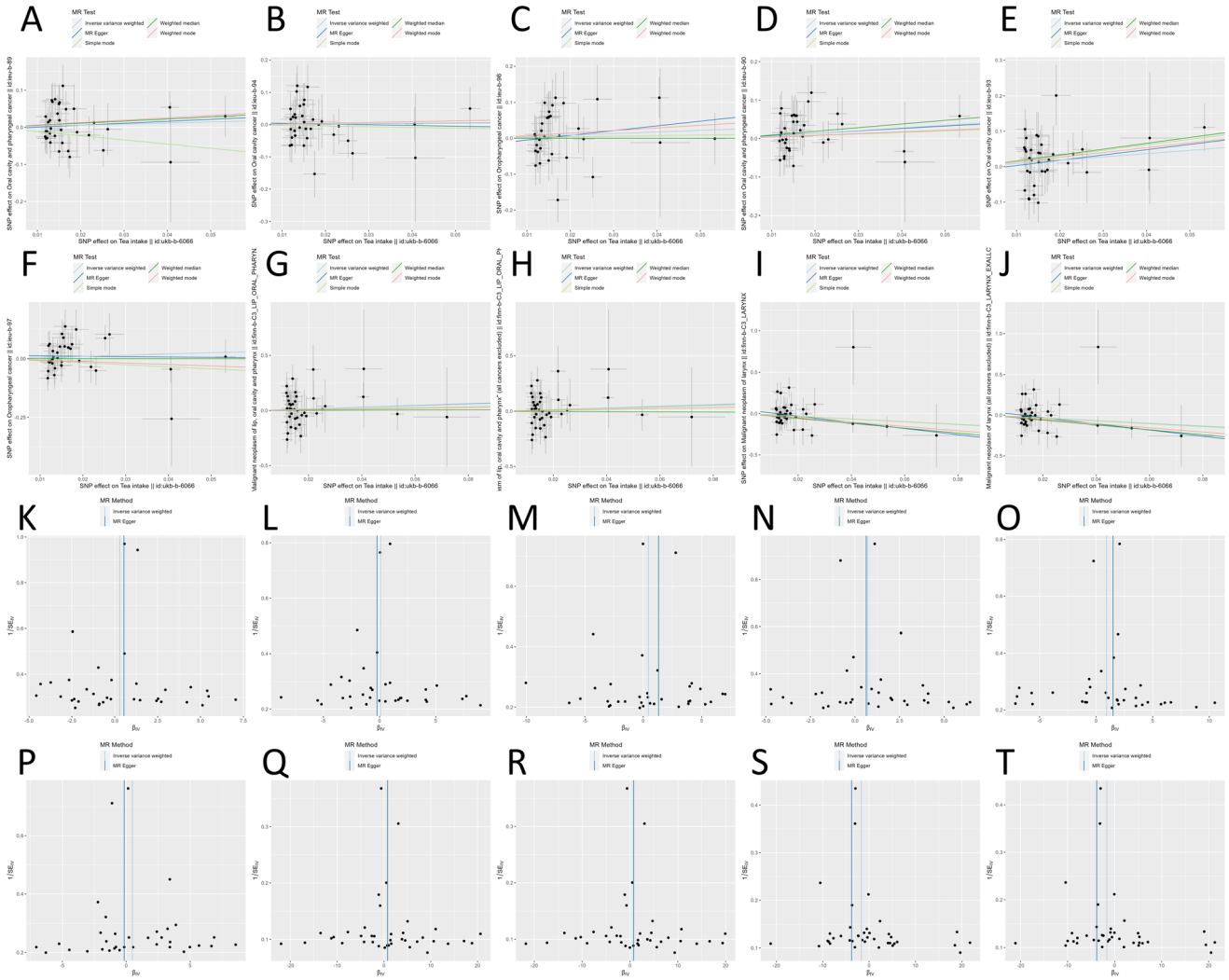

Fig. 3 Scatter plot showing the causal estimate for different MR of tea intake on head and neck cancer and the funnel plot from single SNP analyses for tea intake on head and neck cancer. A–J: forest plot; K–T: funnel plot. A and K: oral cavity and pharyngeal cancer (European region); B and L: oral cavity cancer (European region); C and M: oropharyngeal cancer (European region); D and N: oral cavity and pharyngeal cancer (North America region); E and O: oral cavity cancer (North America region); F and P: oropharyngeal cancer (North America region); G and Q: malignant neoplasm of the lips, oral cavity and pharynx; H and R: malignant neoplasm of the lips, oral cavity and pharynx (all cancers excluded); I and S: malignant neoplasm of the larynx; J and T: malignant neoplasm of the larynx (all cancers excluded).

milk to rooibos tea, the lasso regression results showed that added milk to rooibos tea had a causal effect on oral cavity cancer (North American population of Oncoarray) (OR_{lasso} = 0.005, [range: 0.0001–0.323]); tea consumed || id: ukb-b-17988 also has causal effects on oral cavity and pharyngeal cancer (European population of Oncoarray) ($OR_{MR\text{-}Egger}$ = 49.722, [range: 1.750–1412.860], $p_{MR\text{-}Egger}$ = 0.022), malignant neoplasm of the larynx ($OR_{MR\text{-}Egger}$ = 0.00003, [range: 8.565 × 10^{-9}–0.114], $p_{MR\text{-}Egger}$ = 0.013) and malignant neoplasm of the larynx (all cancers excluded) ($OR_{MR\text{-}Egger}$ = 2.252 × 10^{-5}, [range: 6.389 × 10^{-9}–0.079], $p_{MR\text{-}Egger}$ = 0.010); however, lasso regression does not support this result (Table S3†). However, these results mentioned above did not pass the Bonferroni correction *p*-value level (0.05/3 = 0.017). In conclusion, the results of MVMR showed limited evidence to confirm a causal effect of tea consumption on head and neck cancer.

Sensitivity analysis

No outliers were observed in the MR-Egger or IVW tests, except for decaffeinated tea intake on oral cavity and pharyngeal cancer (European population). After removing outliers from MR-PRESSO, MR-PRESSO did not suggest a causal effect. Horizontal pleiotropy between the instrumental variables and outcomes was assessed using the MR-Egger regression. No evidence of horizontal pleiotropy was found, except for added milk to rooibos tea on oropharyngeal cancer (European population). Due to the presence of multicollinearity, the results of MR Egger showed a causal effect on *p*-values less than 0.05. However, because a relatively lenient *p*-value threshold of 5 × 10^{-6} was used when selecting instrumental variables, it could not be definitively concluded that a causal effect necessarily exists. Especially after undergoing multivariate MR for con-

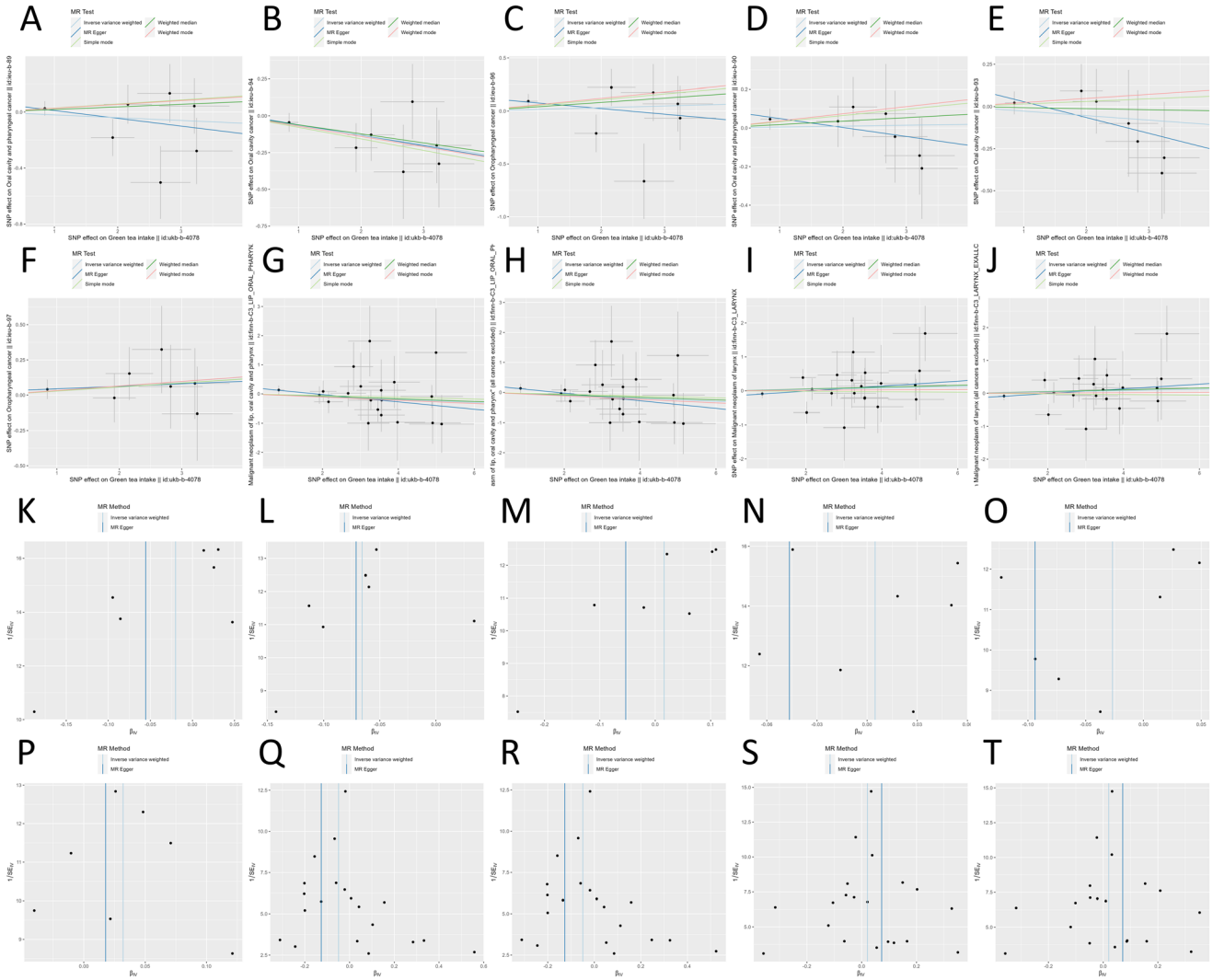

Fig. 4 Scatter plot showing the causal estimate for different MR of green tea intake on head and neck cancer and the funnel plot from single SNP analyses for green tea intake on head and neck cancer. A–J: forest plot; K–T: funnel plot. A and K: oral cavity and pharyngeal cancer (European region); B and L: oral cavity cancer (European region); C and M: oropharyngeal cancer (European region); D and N: oral cavity and pharyngeal cancer (North America region); E and O: oral cavity cancer (North America region); F and P: oropharyngeal cancer (North America region); G and Q: malignant neoplasm of the lips, oral cavity and pharynx; H and R: malignant neoplasm of the lips, oral cavity and pharynx (all cancers excluded); I and S: malignant neoplasm of the larynx; J and T: malignant neoplasm of the larynx (all cancers excluded).

founding factor correction, the effect was no longer statistically significant (Table S4†). Additionally, the leave-one-out analysis revealed no significant difference in the causal estimations of tea intake in the three datasets of head and neck cancer, suggesting that none of the identified causal associations were driven by a single IV. Forest plots of causal effects using a single SNP showed that none were significantly associated with the outcomes (Fig. S1–S26†).

Discussion

In the current study, we conducted MR analyses to assess the causal association between tea intake and head and neck cancer. Utilizing large-scale pooled statistics from tea intake and various head and neck cancer GWAS, the primary MR analysis revealed that green tea consumption served as a protective factor against oral cancer (geographic region: Europe). In secondary MR analysis, adding milk to rooibos tea has a protective effect against malignant neoplasm of the lips, oral cavity, and pharynx. However, after adjusting for smoking and drinking habits and applying the Bonferroni correction, these causal effects were seen to be no longer statistically significant. Our evidence is insufficient to prove that tea intake has a protective effect against head and neck tumors.

Tea has some cancer-suppressing effects,[25] though controversy and uncertainty[26] persist. Studies have demonstrated that tea contains a variety of chemical compounds, such as

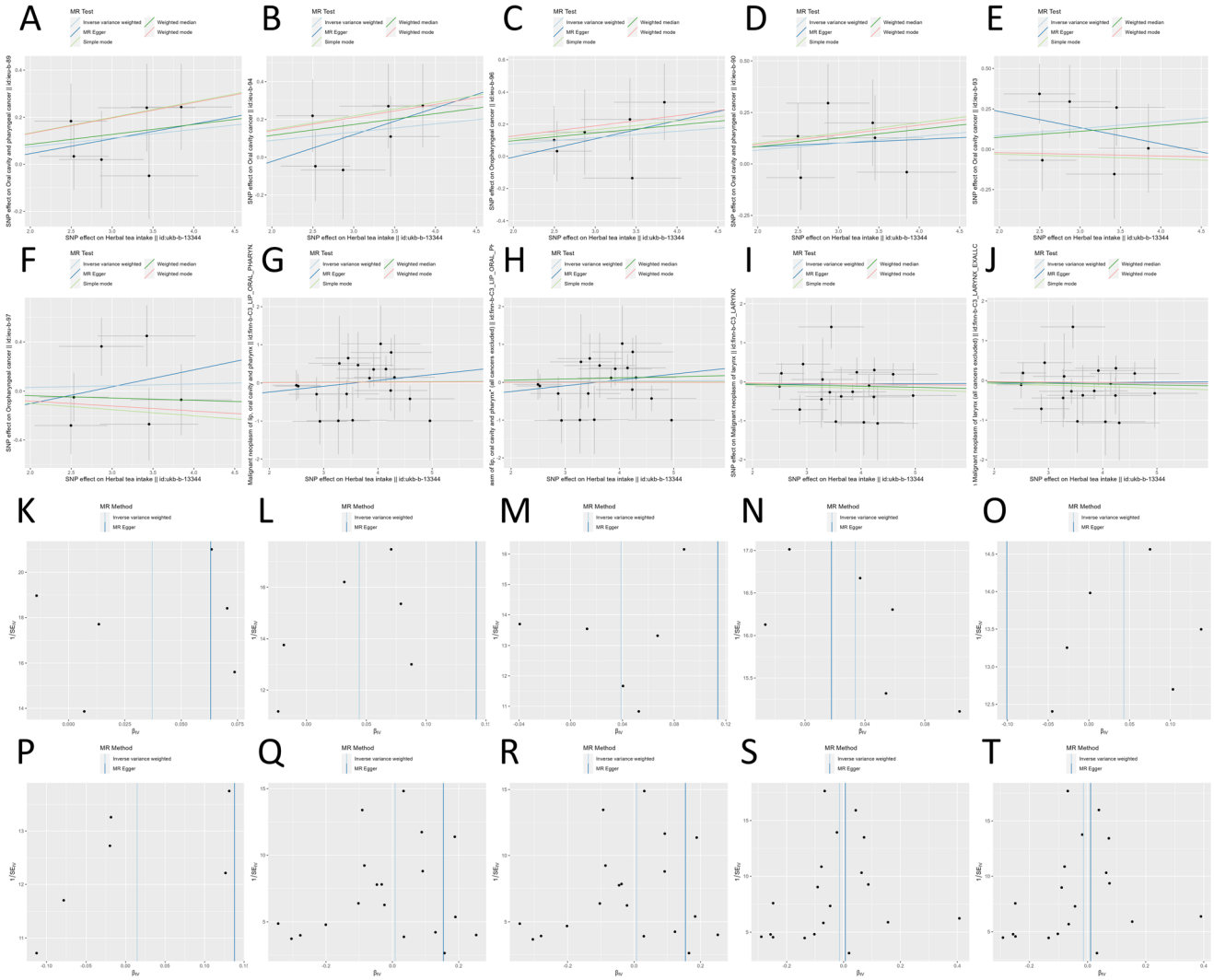

Fig. 5 Scatter plot showing the causal estimate for different MR of herbal tea intake on head and neck cancer and the funnel plot from single SNP analyses for green tea intake on head and neck cancer. A–J: forest plot; K–T: funnel plot. A and K: oral cavity and pharyngeal cancer (European region); B and L: oral cavity cancer (European region); C and M: oropharyngeal cancer (European region); D and N: oral cavity and pharyngeal cancer (North America region); E and O: oral cavity cancer (North America region); F and P: oropharyngeal cancer (North America region); G and Q: malignant neoplasm of the lips, oral cavity and pharynx; H and R: malignant neoplasm of the lips, oral cavity and pharynx (all cancers excluded); I and S: malignant neoplasm of the larynx; J and T: malignant neoplasm of the larynx (all cancers excluded).

epigallocatechin gallate,[27] caffeine, and amino acids, which may exhibit antioxidant,[28] anti-inflammatory, and antitumor effects.[29] Additionally, some tea components may regulate cell proliferation and apoptosis, which may influence the occurrence and development of tumors.[30]

Although some epidemiological studies have indicated that long-term tea consumption may be associated with a reduced risk of certain cancer types, such as gastric,[31] liver,[32–35] and breast cancers,[36–38] other studies have not confirmed the protective effect of tea against specific cancers. Moreover, due to differences such as in the research methods, sources, and tea processing, the cancer-suppressing effect of tea requires further investigation. Consequently, tea cannot be concluded to have a definitive cancer-suppressing effect and should not be regarded as the primary method to prevent or treat cancer.

Green tea has been widely reported to exhibit anticancer effects.[39] Our preliminary results revealed a causal relationship between green tea intake and a reduced risk of oral cancer, whereas those in other anatomical sites (such as the larynx and oropharynx) were not verified. This is consistent with previous studies showing that tea intake has a protective effect on oral cancer, but has not been observed to have the same effect on oropharyngeal cancer.[5] However, this protective effect was not observed after undergoing multivariate MR for confounding factor correction; the effect was no longer statistically significant.

In our multivariate MR with 5×10^{-6} as the standard to extract instrumental variables, our findings also indicated that adding milk to standard tea was a risk factor for malignant neoplasm of the larynx (all cancers excluded). After adjusting

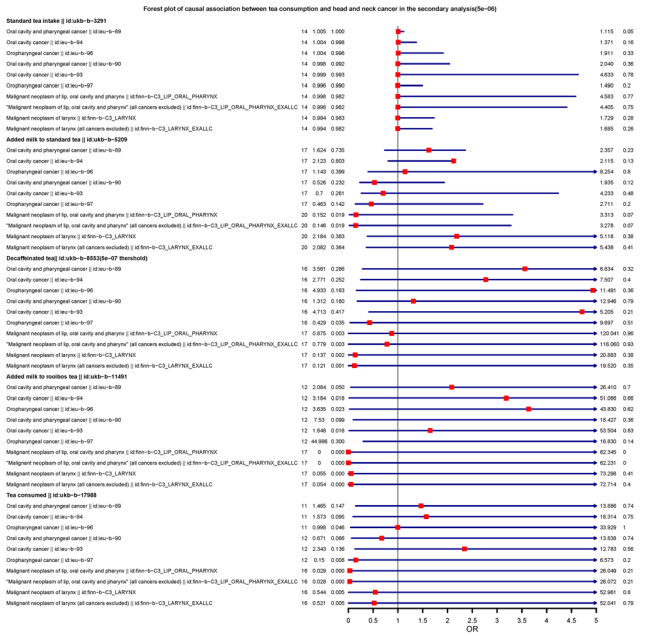

Fig. 6 Forest plot of the causal association between tea consumption and head and neck cancer in the secondary analysis (5×10^{-6}). NSNP, number of SNPs; OR, odds ratio; IVW, inverse variance weighted; CI, confidence interval.

for smoking and drinking habits and applying the Bonferroni correction, this causal effect decreased, suggesting that the effect may not be of practical significance.

We searched the PUBMED database for studies on the association between adding milk to standard tea and tumors, and no studies were found. However, a meta-analysis reported that milk intake has a protective effect against oral and oropharyngeal cancer.[40] Additionally, milk-derived proteins (such as lactoferrin, whey protein, and casein) have been reported to inhibit tumor growth and regulate the expression of cancer-related genes and tumor cytotoxicity.[41]

Our study had several advantages. We analyzed all published GWAS of phenotypes related to tea intake and head and neck tumors from multiple anatomical sites using multiple data sources. Studies have shown that tea consumption can prevent oral cancer in nonsmokers and nondrinkers; however, this effect may be masked in smokers or drinkers.[42] Owing to the strong correlation between smoking, drinking, and tea intake, we also adjusted for the smoking and drinking status to avoid potential confounding factors.[43] The F statistics of all instrumental variables were greater than 10, indicating that there were no weak tool variables. Our conclusion passed the heterogeneity and pleiotropy tests, and we believed that this causal conclusion was more solid.

Although tea has been widely reported for its anticancer and antioxidant effects owing to polyphenols, it should be noted that it needs to be brewed at high temperatures. Although there is insufficient evidence to show that hot drinks can cause head and neck cancer, long-term consumption of hot drinks may damage the esophageal mucosa and increase the risk of esophageal cancer.[44] As an anatomical continuation of the oral cavity and oropharynx, we deemed that attaching importance to the carcinogenic effect of tea as a hot drink, which may mask the anti-cancer effect of substances such as tea polyphenols, is reasonable. As a result, we concluded that there is no causal relationship between tea intake and head and neck cancer.

Our study presented several limitations. First, we could not extract instrumental variables for some exposures under the 5×10^{-8} criteria; thus, we relaxed the threshold to 5×10^{-6}. Consequently, the causal effects may be slightly underestimated. Additionally, the data sourced from the UK Biobank may have some biases in the sample collection process. Some studies suggest the presence of a "healthy volunteer" selection bias in the UK Biobank data.[45] Nevertheless, an effective assessment of exposure–disease relationships can provide broad generalizability and may not require participants to represent the entire population.

Conclusion

In conclusion, we comprehensively examined the potential causal relationship between tea consumption and head and neck cancer. The existing evidence does not support a causal relationship between tea intake and head and neck cancer. Our findings contribute new insights into the risk factors and pathogenesis of head and neck cancer, warranting further verification through additional observational and experimental studies.

Author contributions

Y.X. and X.J., conceptualization; C.D., data curation; M.W. and H.W., formal analysis; X. J. and Y.X., funding acquisition; Y.H., investigation; Q.Z., methodology; X.J., project administration; X.J., resources; Q.Z., software; X.J., supervision; Y.X. and X.J., validation; X.J., visualization; Q.Z. and M.W., roles/writing – original draft; Q.Z., writing – review and editing.

Conflicts of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Acknowledgements

This study was funded by the Jilin Provincial Science and Technology Foundation (grant numbers 20230508064RC and 20210402002GH) and the Achievement Transformation Guiding Foundation of the First Hospital of Jilin University (grant number CGZHYD202012-029). We would like to thank

Editage (https://www.editage.cn) for the English language editing.

References

1 S. Warnakulasuriya, Global epidemiology of oral and oropharyngeal cancer, *Oral Oncol.*, 2009, **45**, 309–316.

2 M. Hashibe, P. Brennan, S.-C. Chuang, S. Boccia, X. Castellsague, C. Chen, M. P. Curado, L. Dal Maso, A. W. Daudt, E. Fabianova, L. Fernandez, V. Wünsch-Filho, S. Franceschi, R. B. Hayes, R. Herrero, K. Kelsey, S. Koifman, C. La Vecchia, P. Lazarus, F. Levi, J. J. Lence, D. Mates, E. Matos, A. Menezes, M. D. McClean, J. Muscat, J. Eluf-Neto, A. F. Olshan, M. Purdue, P. Rudnai, S. M. Schwartz, E. Smith, E. M. Sturgis, N. Szeszenia-Dabrowska, R. Talamini, Q. Wei, D. M. Winn, O. Shangina, A. Pilarska, Z.-F. Zhang, G. Ferro, J. Berthiller and P. Boffetta, Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium, *Cancer Epidemiol. Biomarkers Prev.*, 2009, **18**, 541–550.

3 K. K. Ang, J. Harris, R. Wheeler, R. Weber, D. I. Rosenthal, P. F. Nguyen-Tân, W. H. Westra, C. H. Chung, R. C. Jordan, C. Lu, H. Kim, R. Axelrod, C. C. Silverman, K. P. Redmond and M. L. Gillison, Human papillomavirus and survival of patients with oropharyngeal cancer, *N. Engl. J. Med.*, 2010, **363**, 24–35.

4 J. E. Heck, J. Berthiller, S. Vaccarella, D. M. Winn, E. M. Smith, O. Shan'gina, S. M. Schwartz, M. P. Purdue, A. Pilarska, J. Eluf-Neto, A. Menezes, M. D. McClean, E. Matos, S. Koifman, K. T. Kelsey, R. Herrero, R. B. Hayes, S. Franceschi, V. Wünsch-Filho, L. Fernández, A. W. Daudt, M. P. Curado, C. Chen, X. Castellsagué, G. Ferro, P. Brennan, P. Boffetta and M. Hashibe, Sexual behaviours and the risk of head and neck cancers: a pooled analysis in the International Head and Neck Cancer Epidemiology (INHANCE) consortium, *Int. J. Epidemiol.*, 2010, **39**, 166–181.

5 T. L. Kim, G. H. Jeong, J. W. Yang, K. H. Lee, A. Kronbichler, H. J. van der Vliet, G. Grosso, F. Galvano, D. Aune, J. Y. Kim, N. Veronese, B. Stubbs, M. Solmi, A. Koyanagi, S. H. Hong, E. Dragioti, E. Cho, L. F. M. de Rezende, E. L. Giovannucci, J. I. Shin and G. Gamerith, Tea Consumption and Risk of Cancer: An Umbrella Review and Meta-Analysis of Observational Studies, *Adv. Nutr.*, 2020, **11**, 1437–1452.

6 L.-G. Zhao, Z.-Y. Li, G.-S. Feng, X.-W. Ji, Y.-T. Tan, H.-L. Li, M. J. Gunter and Y.-B. Xiang, Tea Drinking and Risk of Cancer Incidence: A Meta-Analysis of Prospective Cohort Studies and Evidence Evaluation, *Adv. Nutr.*, 2021, **12**, 402–412.

7 B. Elsworth, M. Lyon, T. Alexander, Y. Liu, P. Matthews, J. Hallett, P. Bates, T. Palmer, V. Haberland, G. D. Smith, J. Zheng, P. Haycock, T. R. Gaunt and G. Hemani, The MRC IEU OpenGWAS data infrastructure, *bioRxiv*, 2020, DOI: 10.1101/2020.08.10.244293.

8 C. Lesseur, B. Diergaarde, A. F. Olshan, V. Wünsch-Filho, A. R. Ness, G. Liu, M. Lacko, J. Eluf-Neto, S. Franceschi, P. Lagiou, G. J. Macfarlane, L. Richiardi, S. Boccia, J. Polesel, K. Kjaerheim, D. Zaridze, M. Johansson, A. M. Menezes, M. P. Curado, M. Robinson, W. Ahrens, C. Canova, A. Znaor, X. Castellsagué, D. I. Conway, I. Holcátová, D. Mates, M. Vilensky, C. M. Healy, N. Szeszenia-Dąbrowska, E. Fabiánová, J. Lissowska, J. R. Grandis, M. C. Weissler, E. H. Tajara, F. D. Nunes, M. B. de Carvalho, S. Thomas, R. J. Hung, W. H. M. Peters, R. Herrero, G. Cadoni, H. B. Bueno-de-Mesquita, A. Steffen, A. Agudo, O. Shangina, X. Xiao, V. Gaborieau, A. Chabrier, D. Anantharaman, P. Boffetta, C. I. Amos, J. D. McKay and P. Brennan, Genome-wide association analyses identify new susceptibility loci for oral cavity and pharyngeal cancer, *Nat. Genet.*, 2016, **48**, 1544–1550.

9 K. Burrows, C. J. Bull, T. Dudding, M. Gormley, T. Robinson, V. Tan, J. Yarmolinsky, P. C. Haycock and M. I. E. U. Ieu, *Genome-wide Association Study of Cancer Risk in UK Biobank*, 2021.

10 J.-J. Ni, Q. Xu, S.-S. Yan, B.-X. Han, H. Zhang, X.-T. Wei, G.-J. Feng, M. Zhao, Y.-F. Pei and L. Zhang, Gut Microbiota and Psychiatric Disorders: A Two-Sample Mendelian Randomization Study, *Front. Microbiol.*, 2021, **12**, 737197.

11 S. Burgess, A. Butterworth and S. G. Thompson, Mendelian randomization analysis with multiple genetic variants using summarized data, *Genet. Epidemiol.*, 2013, **37**, 658–665.

12 B. L. Pierce and S. Burgess, Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators, *Am. J. Epidemiol.*, 2013, **178**, 1177–1184.

13 J. Bowden, G. Davey Smith and S. Burgess, Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression, *Int. J. Epidemiol.*, 2015, **44**, 512–525.

14 J. Bowden, G. Davey Smith, P. C. Haycock and S. Burgess, Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator, *Genet. Epidemiol.*, 2016, **40**, 304–314.

15 M. Verbanck, C.-Y. Chen, B. Neale and R. Do, Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases, *Nat. Genet.*, 2018, **50**, 693–698.

16 O. O. Yavorska and S. Burgess, MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data, *Int. J. Epidemiol.*, 2017, **46**, 1734–1739.

17 L. Gaziano, C. Giambartolomei, A. C. Pereira, A. Gaulton, D. C. Posner, S. A. Swanson, Y. L. Ho, S. K. Iyengar, N. M. Kosik, M. Vujkovic, D. R. Gagnon, A. P. Bento, I. Barrio-Hernandez, L. Rönnblom, N. Hagberg, C. Lundtoft, C. Langenberg, M. Pietzner, D. Valentine, S. Gustincich, G. G. Tartaglia, E. Allara, P. Surendran,

S. Burgess, J. H. Zhao, J. E. Peters, B. P. Prins, E. D. Angelantonio, P. Devineni, Y. Shi, K. E. Lynch, S. L. DuVall, H. Garcon, L. O. Thomann, J. J. Zhou, B. R. Gorman, J. E. Huffman, C. J. O'Donnell, P. S. Tsao, J. C. Beckham, S. Pyarajan, S. Muralidhar, G. D. Huang, R. Ramoni, P. Beltrao, J. Danesh, A. M. Hung, K. M. Chang, Y. V. Sun, J. Joseph, A. R. Leach, T. L. Edwards, K. Cho, J. M. Gaziano, A. S. Butterworth and J. P. Casas, Actionable druggable genome-wide Mendelian randomization identifies repurposing opportunities for COVID-19, *Nat. Med.*, 2021, **27**, 668–676.

18 S. Burgess and S. G. Thompson, Interpreting findings from Mendelian randomization using the MR-Egger method, *Eur. J. Epidemiol.*, 2017, **32**, 377–389.

19 M. Verbanck, C. Y. Chen, B. Neale and R. Do, Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases, *Nat. Genet.*, 2018, **50**, 693–698.

20 S. Burgess and S. G. Thompson, Avoiding bias from weak instruments in Mendelian randomization studies, *Int. J. Epidemiol.*, 2011, **40**, 755–764.

21 H. Shim, D. I. Chasman, J. D. Smith, S. Mora, P. M. Ridker, D. A. Nickerson, R. M. Krauss and M. Stephens, A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians, *PLoS One*, 2015, **10**, e0120758.

22 N. Papadimitriou, N. Dimou, K. K. Tsilidis, B. Banbury, R. M. Martin, S. J. Lewis, N. Kazmi, T. M. Robinson, D. Albanes, K. Aleksandrova, S. I. Berndt, D. Timothy Bishop, H. Brenner, D. D. Buchanan, B. Bueno-de-Mesquita, P. T. Campbell, S. Castellví-Bel, A. T. Chan, J. Chang-Claude, M. Ellingjord-Dale, J. C. Figueiredo, S. J. Gallinger, G. G. Giles, E. Giovannucci, S. B. Gruber, A. Gsur, J. Hampe, H. Hampel, S. Harlid, T. A. Harrison, M. Hoffmeister, J. L. Hopper, L. Hsu, J. María Huerta, J. R. Huyghe, M. A. Jenkins, T. O. Keku, T. Kühn, C. La Vecchia, L. Le Marchand, C. I. Li, L. Li, A. Lindblom, N. M. Lindor, B. Lynch, S. D. Markowitz, G. Masala, A. M. May, R. Milne, E. Monninkhof, L. Moreno, V. Moreno, P. A. Newcomb, K. Offit, V. Perduca, P. D. P. Pharoah, E. A. Platz, J. D. Potter, G. Rennert, E. Riboli, M.-J. Sánchez, S. L. Schmit, R. E. Schoen, G. Severi, S. Sieri, M. L. Slattery, M. Song, C. M. Tangen, S. N. Thibodeau, R. C. Travis, A. Trichopoulou, C. M. Ulrich, F. J. B. van Duijnhoven, B. Van Guelpen, P. Vodicka, E. White, A. Wolk, M. O. Woods, A. H. Wu, U. Peters, M. J. Gunter and N. Murphy, Physical activity and risks of breast and colorectal cancer: a Mendelian randomisation analysis, *Nat. Commun.*, 2020, **11**, 597.

23 G. Hemani, J. Zheng, B. Elsworth, K. H. Wade, V. Haberland, D. Baird, C. Laurin, S. Burgess, J. Bowden, R. Langdon, V. Y. Tan, J. Yarmolinsky, H. A. Shihab, N. J. Timpson, D. M. Evans, C. Relton, R. M. Martin, G. Davey Smith, T. R. Gaunt and P. C. Haycock, The MR-Base platform supports systematic causal inference across the human phenome, *Elife*, 2018, **7**, e34408.

24 V. W. Skrivankova, R. C. Richmond, B. A. R. Woolf, N. M. Davies, S. A. Swanson, T. J. VanderWeele, N. J. Timpson, J. P. T. Higgins, N. Dimou, C. Langenberg, E. W. Loder, R. M. Golub, M. Egger, G. Davey Smith and J. B. Richards, Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration, *Br. Med. J.*, 2021, **375**, n2233.

25 C. S. Yang, X. Wang, G. Lu and S. C. Picinich, Cancer prevention by tea: animal studies, molecular mechanisms and human relevance, *Nat. Rev. Cancer*, 2009, **9**, 429–439.

26 M. Eisenstein, Tea's value as a cancer therapy is steeped in uncertainty, *Nature*, 2019, **566**(7742), S6–S7.

27 F. Li, S. Qasim, D. Li and Q. P. Dou, Updated review on green tea polyphenol epigallocatechin-3-gallate as a cancer epigenetic regulator, *Semin. Cancer Biol.*, 2022, **83**, 335–352.

28 I. C. Burckhardt, D. Gozal, E. Dayyat, Y. Cheng, R. C. Li, A. D. Goldbart and B. W. Row, Green tea catechin polyphenols attenuate behavioral and oxidative responses to intermittent hypoxia, *Am. J. Respir. Crit. Care Med.*, 2008, **177**, 1135–1141.

29 E. G. de Mejia, M. V. Ramirez-Mares and S. Puangpraphant, Bioactive components of tea: cancer, inflammation and behavior, *Brain, Behav., Immun.*, 2009, **23**, 721–731.

30 M. Alam, S. Ali, G. M. Ashraf, A. L. Bilgrami, D. K. Yadav and M. I. Hassan, Epigallocatechin 3-gallate: From green tea to cancer therapeutics, *Food Chem.*, 2022, **379**, 132135.

31 X. Li, C. Yu, Y. Guo, Z. Bian, Z. Shen, L. Yang, Y. Chen, Y. Wei, H. Zhang, Z. Qiu, J. Chen, F. Chen, Z. Chen, J. Lv and L. Li, Association between tea consumption and risk of cancer: a prospective cohort study of 0.5 million Chinese adults, *Eur. J. Epidemiol.*, 2019, **34**, 753–763.

32 M. Shimizu, Y. Shirakami, H. Sakai, M. Kubota, T. Kochi, T. Ideta, T. Miyazaki and H. Moriwaki, Chemopreventive potential of green tea catechins in hepatocellular carcinoma, *Int. J. Mol. Sci.*, 2015, **16**, 6124–6139.

33 J. D. Yang and H. Malhi, Green tea consumption: A potential chemopreventive measure for hepatocellular carcinoma?, *Hepatology*, 2018, **67**, 10–12.

34 Y. Li, S.-C. Chang, B. Y. Goldstein, W. L. Scheider, L. Cai, N.-C. Y. You, H. P. Tarleton, B. Ding, J. Zhao, M. Wu, Q. Jiang, S. Yu, J. Rao, Q.-Y. Lu, Z.-F. Zhang and L. Mu, Green tea consumption, inflammation and the risk of primary hepatocellular carcinoma in a Chinese population, *Cancer Epidemiol.*, 2011, **35**, 362–368.

35 C. Bamia, P. Lagiou, M. Jenab, A. Trichopoulou, V. Fedirko, K. Aleksandrova, T. Pischon, K. Overvad, A. Olsen, A. Tjønneland, M.-C. Boutron-Ruault, G. Fagherazzi, A. Racine, T. Kuhn, H. Boeing, A. Floegel, V. Benetou, D. Palli, S. Grioni, S. Panico, R. Tumino, P. Vineis, H. B. Bueno-de-Mesquita, V. K. Dik, N. Bhoo-Pathy, C. S. P. M. Uiterwaal, E. Weiderpass, E. Lund, J. R. Quirós, R. Zamora-Ros, E. Molina-Montes, M.-D. Chirlaque, E. Ardanaz, M. Dorronsoro, B. Lindkvist, P. Wallström, L. M. Nilsson, M. Sund, K.-T. Khaw, N. Wareham, K. E. Bradbury, R. C. Travis, P. Ferrari, T. Duarte-Salles,

M. Stepien, M. Gunter, N. Murphy, E. Riboli and D. Trichopoulos, Coffee, tea and decaffeinated coffee in relation to hepatocellular carcinoma in a European population: multicentre, prospective cohort study, *Int. J. Cancer*, 2015, **136**, 1899–1908.

36 Y. Wang, Y. Zhao, F. Chong, M. Song, Q. Sun, T. Li, L. Xu and C. Song, A dose-response meta-analysis of green tea consumption and breast cancer risk, *Int. J. Food Sci. Nutr.*, 2020, **71**, 656–667.

37 S. Yu, L. Zhu, K. Wang, Y. Yan, J. He and Y. Ren, Green tea consumption and risk of breast cancer: A systematic review and updated meta-analysis of case-control studies, *Medicine*, 2019, **98**, e16147.

38 V. Gianfredi, D. Nucci, A. Abalsamo, M. Acito, M. Villarini, M. Moretti and S. Realdon, Green Tea Consumption and Risk of Breast Cancer and Recurrence-A Systematic Review and Meta-Analysis of Observational Studies, *Nutrients*, 2018, **10**(12), 1886.

39 M. S. Butt and M. T. Sultan, Green tea: nature's defense against malignancies, *Crit. Rev. Food Sci. Nutr.*, 2009, **49**, 463–473.

40 J. Yuan, W. Li, W. Sun and S. Deng, Milk and dairy products consumption and the risk of oral or oropharyngeal cancer: a meta-analysis, *Biosci. Rep.*, 2019, **39**(12), BSR20193526.

41 T. Wang, X. Liu, Y. Y. Ng, K. Tarleton, A. Tran, T. Tran, W. Y. Xue, P. Youssef, P. Yuan, D. Zhang, R. Paolini and A. Celentano, Milk-Derived Proteins and Peptides in Head and Neck Carcinoma Treatment, *Biomolecules*, 2022, **12**(2), 290.

42 F. Chen, B. C. He, L. J. Yan, F. P. Liu, J. F. Huang, Z. J. Hu, Z. Lin, X. Y. Zheng, L. S. Lin, Z. F. Zhang and L. Cai, Tea consumption and its interactions with tobacco smoking and alcohol drinking on oral cancer in southeast China, *Eur. J. Clin. Nutr.*, 2017, **71**, 481–485.

43 W. Dongmeng, X. Yu'e, G. Wenjing, Z. Ke, L. Jun, Y. Canqing, W. Shengfeng, H. Tao, S. Dianjianyi, L. Chunxiao, P. Yuanjie, P. Zengchang, Y. Min, W. Hua, W. Xianping, D. Zhong, W. Fan, J. Guohong, W. Xiaojie, L. Yu, D. Jian, L. Lin, C. Weihua and L. Liming, Heritability of tea drinking and its relationship with cigarette smoking in the Chinese male adult twins, *Addict. Biol.*, 2022, **27**, e13129.

44 C. Yu, H. Tang, Y. Guo, Z. Bian, L. Yang, Y. Chen, A. Tang, X. Zhou, X. Yang, J. Chen, Z. Chen, J. Lv and L. Li, 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**, 489–497.

45 A. Fry, T. J. Littlejohns, C. Sudlow, N. Doherty, L. Adamska, T. Sprosen, R. Collins and N. E. Allen, Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population, *Am. J. Epidemiol.*, 2017, **186**, 1026–1034.