nature portfolio

Peer Review File

Smoking and Alcohol by HPV Status in Head and Neck Cancer A Mendelian Randomization Study

Editorial note: Parts of this Peer Review File have been redacted as indicated to remove third-party material where no permission to publish could be obtained.

REVIEWER COMMENTS

Reviewer #1 - Head and neck cancer, HPV (Remarks to the Author):
I am not an epidemiologist and am not qualified to comment on the approach, methodology or statistical analysis, but have been asked to give my assessment of the HPV aspects of this work. My critique is therefore limited to the significance and originality of the findings.

HPV-associated HNSCC is a growing clinical problem, with the current trend of rising incidence in the USA expected to continue until 2045, even assuming high gender neutral HPV vaccine coverage. The manuscript addresses the important question as to whether tobacco smoking and alcohol consumption (established risk factors for HPV-negative HNSCC) may also play a causal role in HPV-positive HNSCC. Strengths include the large cohort size and potentially the application of Mendelian randomization to overcome the issue of unmeasured confounding, although as stated above, I leave it to other reviewers to provide expert comment on the methods applied.

HPV status was primarily assigned based on the detection of anti-E6 or anti-E7 antibodies in participant serum samples - an indirect but robust approach that has been validated, including in previous studies by these authors. In cases where antibody status was unknown, the authors were able to obtain p16INK4A staining data from tumour tissue - a highly sensitive and specific HPV biomarker in SCC of the oropharynx.

I would recommend modifying the phrase "HPV p16 immunohistochemistry result", since this misleadingly suggests p16 is a viral protein, whereas it is in fact a host protein that displays elevated expression in cells with high E7 expression.

Although at least 90% of HPV-positive OPSCC cases harbour HPV16, a significant minority are caused by other HPV types, including HPV31 and HPV33. I think the authors should

acknowledge this (as they also acknowledge that as many as 4-5% of their oral SCC cases are likely HPV-positive) and if they could comment on whether their assay to detect HPV E6/E7-specific antibodies would also be expected to detect antibodies against E6/E7 from these minority HPV types. If not, then together with discounting the HPV-positive oral SCC cases, failure to detect non-HPV16 types (at least in cases where p16 IHC was not used) is another reason for a likely underestimation of HPV-positivity in the cohort.

Finally, for publication in a broad interest journal such as Nature Communications, I would recommend the authors provide a clearer explanation of their MR approach and why it they have employed it to make their study more accessible to the readership.

Reviewer #2 - Head and neck cancer, epidemiology (Remarks to the Author):
Important question. Robust study designs from clinical epidemiological perspective.
Would suggest explaining the methods better both in the manuscript and the tables. The interpretations are based upon methods that are not necessarily broadly understood and hard for me to understand beyond direction/ magnitude of change, CIs. So needs to be more interpretable to the regular person.
Helpful to understand influence of tobacco in a way not previously understood.

Reviewer #3 - Mendelian randomisation (Remarks to the Author):
Differences in Causal Effects of Smoking and Alcohol by HPV Status in Head and Neck Cancer: A Multivariable Mendelian Randomization Study

By Thakral et al

The authors investigated the causal effect of alcohol and smoking on HPV-positive and HPV-negative head and neck squamous cell carcinoma (HNSCC). One and two sample analyses were conducted using published data and data generated by the authors. The results reveal that smoking and alcohol independently increase the risk of both HPV-positive and HPV-negative HNSCC with similar causal effects.

Overall, this is an interesting project that aims to understand if HPV-positive HNSCC shares the same risk factors as the more common HPV-negative HNSCC. The conclusions are clearly supported by the evidence available but there are points that can benefit from further clarification. In more detail:

1) My biggest concern is what the results really mean. The results suggest almost no difference in the effect of smoking and alcohol between HPV-positive and HPV-negative HNSCC cases. This can be extended to include causal logic, in which case the question is: are HPV-positive cases just cases with HPV or cases due to HPV? Before we can understand the causal contribution of smoking and alcohol on HPV strata, we first need to understand the causal contribution of the HPV status in the two strata.
2) In the definition of HPV-positive and HPV-negative HNSCC cases there are a number of assumptions made. Do these assumptions affect the results of the study?
3) Recently the use of a negative control has been suggested for MR studies (DOI: 10.1093/ije/dyaa288). Can the authors include such as test in this case?
4) Alcohol is one of the behavioural traits strongly associated with non-linear effects (doi: 10.1371/journal.pmed.1003178). Have the authors tested for the presence of such differences?
5) There are very marked differences in HPV-positive and HPV-negative HNSCC cases between men and women (STable 1). These will also be present in smoking and alcohol consumption. How were possible different causal effects between males and females investigated?
6) The Illumina Onco Array is focusing on variants relevant with common cancers (doi: 10.1158/1055-9965.EPI-16-0106). This means that it is not optimal for following up associations of other traits, which is the aim of MR. This may have affected the results presented but it has not been mentioned, or its impact tested.
7) How was ancestry information used in the GWAS? Was it adjusted by PCs or stratified and meta-analysed? I understand that the former was used but all recent large GWASs follow the later or a mixed models approach.
8) Did the GWAS for HPV-positive and HPV-negative HNSCC used the same controls? Should the controls be matched for HPV status?
9) Why was the MVMR ridge regression used?

10) In the paragraph of the Methods starting with “The study population consisted” further down is written “a total of totaling”

11) I am not sure what the statement “clumping significance levels for index and secondary SNPs of 1” means?

12) I am also not sure what this means “Specifically, pairwise associations between SNP associations were set to zero as the samples used for the summary statistics for the two exposures were independent”

13) ‘the “mr function” from’ could be more clearly written as ‘the “mr()” function from’ or ‘the “mr” function from’

14) “we first built the polygenic risk scores (PRS) of the outcomes” Outcomes or exposures? Did the authors mean genetic risk scores (GRSs) as opposed to PRSs? If they meant PRSs, then what method did they use to build PRSs?

15) “we conducted Q statistic minimization yielding a Q-statistic of 188.20 (P=0.36)” What is the significance/interpretation of this?

16) I would put the risk tolerance and high-risk sexual behaviour results as a separate paragraph.

17) Table 1 does not provide the numbers of independent SNPs included as instrument variables for each smoking and alcohol use behaviour

RESPONSE TO REFEREES LETTER
REVIEWER COMMENTS

Reviewer #1 - Head and neck cancer, HPV (Remarks to the Author):

I am not an epidemiologist and am not qualified to comment on the approach, methodology or statistical analysis, but have been asked to give my assessment of the HPV aspects of this work. My critique is therefore limited to the significance and originality of the findings.

HPV-associated HNSCC is a growing clinical problem, with the current trend of rising incidence in the USA expected to continue until 2045, even assuming high gender neutral HPV vaccine coverage. The manuscript addresses the important question as to whether tobacco smoking and alcohol consumption (established risk factors for HPV-negative HNSCC) may also play a causal role in HPV-positive HNSCC. Strengths include the large cohort size and potentially the application of Mendelian randomization to overcome the issue of unmeasured confounding, although as stated above, I leave it to other reviewers to provide expert comment on the methods applied.

HPV status was primarily assigned based on the detection of anti-E6 or anti-E7 antibodies in participant serum samples - an indirect but robust approach that has been validated, including in previous studies by these authors. In cases where antibody status was unknown, the authors were able to obtain p16INK4A staining data from tumour tissue - a highly sensitive and specific HPV biomarker in SCC of the oropharynx.

1. I would recommend modifying the phrase "HPV p16 immunohistochemistry result", since this misleadingly suggests p16 is a viral protein, whereas it is in fact a host protein that displays elevated expression in cells with high E7 expression.

We thank the reviewer recommendation, thus, in the main text of methods. “HPV p16 immunohistochemistry results were utilized if the antibody status was unknown” was replaced with:

“For cases where the HPV16 antibody status was indeterminable, we utilized the expression of p16 as a surrogate marker. p16 is a cellular protein whose overexpression is an indirect measure of HPV-associated oncogenic activity, rather than a direct viral marker.”

2. Although at least 90% of HPV-positive OPSCC cases harbour HPV16, a significant minority are caused by other HPV types, including HPV31 and HPV33. I think the authors should acknowledge this (as they also acknowledge that as many as 4-5% of their oral SCC cases are likely HPV-positive) and if they could comment on whether their assay to detect HPV E6/E7-specific antibodies would also be expected to detect antibodies against E6/E7 from these minority HPV types. If not, then together with discounting the HPV-positive oral SCC cases, failure to detect non-HPV16 types (at least in cases where p16 IHC was not used) is another reason for a likely underestimation of HPV-positivity in the cohort.

The main text of methods have been revised to reflect these valuable points:

“HPV-positive cancers were defined as oropharyngeal cancer (OPC) patients with positive HPV16 antibody status as the primary classifier, given that up to 90% of HPV-positive OPCs are attributed to HPV type 16 (HPV16). OPC cases were classified as HPV-positive or HPV-negative based on a previously validated and described HPV16 seropattern algorithm, acknowledging that this method may

lead to an underrepresentation of seropositivity for other high-risk HPV subtypes such as HPV 18, 31, and 33."

3. Finally, for publication in a broad interest journal such as Nature Communications, I would recommend the authors provide a clearer explanation of their MR approach and why it they have employed it to make their study more accessible to the readership.

We appreciate and agree with the reviewer's suggestion on making the manuscript more accessible to the general readers. We have modified the introduction, methods and discussion to provide additional explanation on the aspects of the Mendelian randomization approach and its uses and advantages.

Added to introduction section:

"Furthermore, multivariable MR allows for the simultaneous estimation of the independent and joint effects of two or more exposures on an outcome, which is particularly relevant given that the combined exposure to tobacco and alcohol has been demonstrated to exert a significant synergistic effect on the incidence of HNSCC."

Additional information has been provided in the method section:

"Mendelian randomization is a statistical method that combines quantified estimates from associations between genetic instruments and the risk factor with parallel associations between these genetic variants and the outcome, to determine an estimate of the risk factor's impact on the outcome."

"The application of multiple MR methods allows the assessment of causal effects across various statistical assumptions, thereby adjusting for potential pleiotropy and invalid instruments."

"Pleiotropy arises when a genetic variant is linked to the outcome of interest through multiple pathways, which may not necessarily involve the exposure under investigation. The presence of pleiotropy can alter both the magnitude and direction of the association between the exposure and outcome. To evaluate whether the assumptions of MR hold true, multiple MR methodologies are employed to assess the consistency of findings across different approaches."

Added to discussion section:

"To address the correlation between smoking and alcohol consumption, as well as to simultaneously explore their independent effects, we performed multivariable MR analyses on HPV-positive and negative HNSCC groups separately. Multivariable MR extends the basic MR framework to accommodate the complexity of multiple correlated exposures, enabling the evaluation of the independent causal effects of smoking and alcohol use on HNSCC risk."

Reviewer #2 - Head and neck cancer, epidemiology (Remarks to the Author):

Important question. Robust study designs from clinical epidemiological perspective.

1. Would suggest explaining the methods better both in the manuscript and the tables. The interpretations are based upon methods that are not necessarily broadly understood and hard for me to understand beyond direction/ magnitude of change, CIs. So needs to be more interpretable to the regular person.

Thank you for this salient feedback. We have further explained the MR approach throughout the introduction, methods and discussion sections to make the study more accessible to the general readership.

Added to introduction section:

“Furthermore, multivariable MR allows for the simultaneous estimation of the independent and joint effects of two or more exposures on an outcome, which is particularly relevant given that the combined exposure to tobacco and alcohol has been demonstrated to exert a significant synergistic effect on the incidence of HNSCC.”

Additional information has been provided in the method section:

“Mendelian randomization is a statistical method that combines quantified estimates from associations between genetic instruments and the risk factor with parallel associations between these genetic variants and the outcome, to determine an estimate of the risk factor's impact on the outcome.”

“The application of multiple MR methods allows the assessment of causal effects across various statistical assumptions, thereby adjusting for potential pleiotropy and invalid instruments.”

“Pleiotropy arises when a genetic variant is linked to the outcome of interest through multiple pathways, which may not necessarily involve the exposure under investigation. The presence of pleiotropy can alter both the magnitude and direction of the association between the exposure and outcome. To evaluate whether the assumptions of MR hold true, multiple MR methodologies are employed to assess the consistency of findings across different approaches.”

Added to discussion section:

“To address the correlation between smoking and alcohol consumption, as well as to simultaneously explore their independent effects, we performed multivariable MR analyses on HPV-positive and negative HNSCC groups separately. Multivariable MR extends the basic MR framework to accommodate the complexity of multiple correlated exposures, enabling the evaluation of the independent causal effects of smoking and alcohol use on HNSCC risk.”

Reviewer #3

The authors investigated the causal effect of alcohol and smoking on HPV-positive and HPV-negative head and neck squamous cell carcinoma (HNSCC). One and two sample analyses were conducted using published data and data generated by the authors. The results reveal that smoking and alcohol independently increase the risk of both HPV-positive and HPV-negative HNSCC with similar causal effects.

Overall, this is an interesting project that aims to understand if HPV-positive HNSCC shares the same risk factors as the more common HPV-negative HNSCC. The conclusions are clearly supported by the evidence available but there are points that can benefit from further clarification. In more detail:

1) My biggest concern is what the results really mean. The results suggest almost no difference in the effect of smoking and alcohol between HPV-positive and HPV-negative HNSCC cases. This can be extended to include causal logic, in which case the question is: are HPV-positive cases just cases with HPV or cases due to HPV? Before we can understand the causal contribution of smoking and alcohol on HPV strata, we first need to understand the causal contribution of the HPV status in the two strata.

The reviewer highlights an important observation regarding the need to discern the causal contributions of HPV infection in HNSCC cases. Our results do indicate that both smoking and alcohol independently increase the risk of both HPV-positive and HPV-negative HNSCC, with lifetime smoking having similar effects and alcohol having a stronger effect in HPV-negative compared to HPV-positive cases. The current understanding in the literature posits that HPV-positive oropharyngeal cancers have a distinct etiopathogenesis compared to their HPV-negative counterparts, often with less pronounced associations with smoking and alcohol use. In our study, the apparent lack of disparity in risk factor impact could suggest a more nuanced interaction between HPV status and these carcinogens than previously understood.

We acknowledge that our study design primarily addresses associations rather than causal mechanisms. To determine causality, further research is required. We propose that our findings serve as a preliminary indication of the complex relationship between HPV, smoking, and alcohol, meriting more in-depth investigation. Mechanistic studies that explore the biological interaction between HPV oncogenes and carcinogen-induced DNA damage in epithelial cells could provide further clarity.

This paragraph has been added to the discussion section:

“Finally, it is important to note that while MR approaches can suggest potential causal relationships, additional evidence is required to confirm causal mechanisms. HPV-positive oropharyngeal cancers are considered to have a distinct etiopathogenesis compared to their HPV-negative counterparts, often with less pronounced associations with smoking and alcohol use. In our study, the apparent lack of disparity in the impact of risk factors could suggest a more complex interaction between HPV status and these carcinogens than previously understood. Mechanistic studies that explore the biological interactions between HPV oncogenes and carcinogen-induced DNA damage in epithelial cells could provide further clarity.”

2) In the definition of HPV-positive and HPV-negative HNSCC cases there are a number of assumptions made. Do these assumptions affect the results of the study?

We appreciate the reviewer highlights on the potential impact of the assumptions made for the determination of HPV status. We explained these assumptions more clearly in the revised methods section while discussing the potential limitations.

Addition to the methods section:

“For cases where the HPV16 antibody status was indeterminable, we utilized the expression of p16 as a surrogate marker. p16 is a cellular protein whose overexpression is an indirect measure of HPV-associated oncogenic activity, rather than a direct viral marker.”

“HPV-positive cancers were defined as oropharyngeal cancer (OPC) patients with positive HPV16 antibody status as the primary classifier, given that up to 90% of HPV-positive OPCs are attributed to HPV type 16 (HPV16). OPC cases were classified as HPV-positive or HPV-negative based on a previously validated and described HPV16 seropattern algorithm, acknowledging that this method may lead to an underrepresentation of seropositivity for other high-risk HPV subtypes such as HPV 18, 31, and 33.”

3) Recently the use of a negative control has been suggested for MR studies (DOI: 10.1093/ije/dyaa288). Can the authors include such as test in this case?

We appreciate the potential benefit of utilizing a negative control as an adjunctive approach to our MR analyses. The negative control method involves the use of variables that are predetermined relative to the phenotypes of interest but are likely to be susceptible to population stratification as negative control outcomes within an MR analysis to detect population stratification biases. This method is only able to detect biases as far as it affects the chosen negative control outcome. Consequently, the absence of effects on negative control outcomes does not conclusively indicate that the phenotype or its associated MR analysis is free from bias. This limitation can be alleviated by selecting negative control outcomes that are more likely to be influenced by population stratification, provided such outcomes are available.

A limitation in our data is the absence of phenotypes suitable as negative controls. We acknowledged that in our study, PCA might not be sufficiently powered to account for biases arising from population stratification. We recommend incorporating negative controls in future research directions to enhance bias detection.

4) Alcohol is one of the behavioural traits strongly associated with non-linear effects (doi: 10.1371/journal.pmed.1003178). Have the authors tested for the presence of such differences?

Alcohol consumption has been shown to have non-linear associations with colorectal cancer with low doses being protective and high doses being a risk factor. Some studies have shown that alcohol increased the risk only at high doses in the absence of tobacco. However, meta-analysis of observational studies have shown that alcohol consumption has a dose-response relationship with HNSCC. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4453639/)

As such, we did not investigate non-linear effects of alcohol on head and neck cancer risk.

[**figure redacted**]

5) There are very marked differences in HPV-positive and HPV-negative HNSCC cases between men and women (STable 1). These will also be present in smoking and alcohol consumption. How were possible different causal effects between males and females investigated?

We appreciate and agree with the reviewer noting the marked differences in alcohol and smoking use by sex/gender. Sex-specific methods such as MrGxE have been described. However, although sex-specific summary data on instrument–outcome could be generated, sex-specific summary data on instrument–exposure associations were not publicly available. The lack of sex-specific instruments excludes the comparisons of effects across sexes and prevents a thorough analysis of potential differential effects between males and females. We agree that this would be an important analysis to explored and will include this as a limitation in our study.

Added to the discussion section:

"Furthermore, the lack of sex-specific instrument exposure information prevented the assessment of smoking and alcohol use stratified by sex, which is particularly relevant given the differences in exposures seen across males and females."

6) The Illumina Onco Array is focusing on variants relevant with common cancers (doi: 10.1158/1055-9965.EPI-16-0106). This means that it is not optimal for following up associations of other traits, which is the aim of MR. This may have affected the results presented but it has not been mentioned, or its impact tested.

We thank the reviewer for making this point. The summary estimates for the exposures were obtained from the GSCAN meta-GWAS which employed multiple different SNP array panels across the various GWAS utilized. The Illlumina OncoArray was used to conduct the GWASs to obtain the estimates for the association of the same SNPs with the outcome (HNSCC). In the two sample MR approach, if the SNPs in exposure GWAS (smoking, alcohol, etc, which are not cancer-associated) were not found in the outcome GWAS (ie OncoArray). then a proxy would be used from the LD panel as we performed in the analysis. We refer to this in our methods section: "The proxies for SNPs missing from the outcome GWAS were generated using the LD proxy tool (European superpopulation reference) and r 2 cut-off of 0.8 ;".

7) How was ancestry information used in the GWAS? Was it adjusted by PCs or stratified and meta-analysed? I understand that the former was used but all recent large GWASs follow the later or a mixed models approach.

We recognize and agree with the reviewer's comment regarding the use and potential benefit of different approaches to adjust for population stratification. Bayesian mixed models approaches have shown an increase in power for GWASs and are increasingly being used in large GWASs. However, "the power gain that mixed models offer over existing methods via their more flexible prior on SNP effect sizes is contingent on the true genetic architecture being sufficiently non-infinitesimal and the sample size being sufficiently large." (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4342297/) This power gain is better with sample size above 30,000. With our sample size being <30,000, we did not expect it to improve the power of our GWASs.

8) Did the GWAS for HPV-positive and HPV-negative HNSCC used the same controls? Should the controls be matched for HPV status?

The controls were the same for both analyses. We classified the cases using HPV status, as these two cancer subtypes have distinct biological and pathophysiological etiologies. Matching the controls for HPV status would be necessary to address the specific associations of HPV with HNSCC, which was not the aim of our study question.

9) Why was the MVMR ridge regression used?

We thank the reviewer for raising this question. The correlation between smoking and alcohol consumption behaviors may lead to collinearity, potentially introducing bias and causing inflated standard errors in multivariable Mendelian randomization (MR) models. To assess the impact of collinearity, we employed ridge regression models that utilize L2 regularization. Although this regularization method introduces a small degree of bias into the regression estimates, it improves the precision of standard errors. Accurate standard errors are important for assessing robust causal inference, especially when analyzing the correlation between two significant risk factors.

10) In the paragraph of the Methods starting with "The study population consisted" further down is written "a total of totaling"

This has been corrected to "Within VOYAGER, OncoArray data were available from a total of 3431 cases and 3469 controls from Europe and North America".

11) I am not sure what the statement "clumping significance levels for index and secondary SNPs of 1" means?

Thank you for noting this phrase requiring clarification. The clumping algorithm that we used sequentially chooses the top SNP, removes all SNPs in LD above some threshold within the r2 window, then goes on to the next top hit and repeats the pruning process until no more SNPs are left above the specified p-value threshold. With clumping significance levels of 1, there is no threshold, and all SNPs in LD with the top SNP are removed.

We have now clarified this statement as follows:

"Using a reference population of 1000 genomes' European superpopulation, all SNPs with p-values < 5x10^-8 were selected as potential index SNPs and then pruned using a clumping window of 10,000 base pairs (bp), with an r^2 cut-off of 0.001 to ensure independence. Secondary SNPs in LD were removed at a threshold levels set at P-values < 1."

12) I am also not sure what this means "Specifically, pairwise associations between SNP associations were set to zero as the samples used for the summary statistics for the two exposures were independent"

The MVMR approach requires pairwise covariances between an instrument and pairs of exposures to be known across all SNPs for testing and sensitivity analyses. If gene-exposure associations are estimated in separate non-overlapping samples, then the covariances will be zero by design. This sentence has been changed to state that the covariances for pairwise associations between SNP associations were zero by design:

We have now clarified this sentence as follows:
"Specifically, the covariances for pairwise associations between SNP-exposure effects were assumed to be zero as the summary statistics for each exposure were derived from independent, non-overlapping samples."

13) 'the "mr function" from' could be more clearly written as 'the "mr()" function from' or 'the "mr" function from'

Thank you for pointing this out, we have revised the sentence to “the “mr()” function from” in the methods section.

14) “we first built the polygenic risk scores (PRS) of the outcomes” Outcomes or exposures? Did the authors mean genetic risk scores (GRSs) as opposed to PRSs? If they meant PRSs, then what method did they use to build PRSs?

We apologize for the confusing expression in our manuscript. The phrase 'we first built the polygenic risk scores (PRS) of the outcomes' should have been more clearly expressed. In our study, the polygenic risk scores (PRS), which is equivalent to genetic risk scores (GRS) in this context, were constructed for each exposure (including smoking and drinking behaviors). We then conducted a one-sample Mendelian Randomization using the PRS of each exposure to assess its effect on the binary outcome of head and neck squamous cell carcinoma.

In detail, the polygenic scores were constructed using the same SNPs as those in our main two-sample MR analysis, applying the same significance threshold and clumping strategy. The weight assigned to each SNP was its effect size (β) extracted from the respective GWAS. We used a weighted sum method to calculate the genetic score for each participant based on these SNPs.

To improve clarity, we have revised the sentence to read: “For factorial MR, we first constructed the polygenic risk scores (PRS) for the exposures as instrumental variables for each study participant.”

15) “we conducted Q statistic minimization yielding a Q-statistic of 188.20 (P=0.36)” What is the significance/interpretation of this?

Weak instruments can increase the false positive rate for pleiotropy detection, as heterogeneity in effect estimates due to weak instrument bias is conflated with heterogeneity as a result of pleiotropic bias. As a correction it is possible to estimate heterogeneity from pleiotropy through Q-statistic minimisation. Where the MVMR assumptions are potentially violated, specifically where instruments are weak or exhibit pleiotropy, it is possible to obtain more robust estimates through Q-statistic minimisation. We conducted Q statistic minimization yielding a Q-statistic of 188.20 (P=0.36), implying a lack of heterogeneity after correction for weak instrument bias.

We have added this interpretation to the revised manuscript as:

In the methods section:

“Whenever assumptions were violated, we obtained more robust estimates through Q-statistic minimization, which is particularly effective when instruments are weak or exhibit pleiotropy.”

In the results section:

“Owing to the weak instrument strengths (<10), we conducted Q statistic minimization yielding a Q-statistic of 188.20 (P=0.36), implying a lack of heterogeneity after correction for weak instrument bias.”

16) I would put the risk tolerance and high-risk sexual behaviour results as a separate paragraph.

The results of this analysis have been revised into a separate paragraph in the results section.

17) Table 1 does not provide the numbers of independent SNPs included as instrument variables for each smoking and alcohol use behaviour.

Please refer to the "SNPs, *N* =" available below the OR in the first row of Table 1.

Table 1. Univariable Mendelian Randomisation of Smoking and Alcohol Consumption Exposures on Oral and Oropharyngeal Cancer Stratified by HPV status.

		Smoking Initiation		Comprehensive Smoking Index		Drinks Per Week	
HPV status	**Method**	**OR (95% CI)*** SNPs, *N* = 57	***P* value**	**OR (95% CI)** SNPs, *N* = 90	***P* value**	**OR (95% CI)** SNPs, *N* = 25	***P* value**
HPV-negative HNSCC vs controls	MR Egger	29.5 (3.46, 251)	3.11E-03	2.46 (0.19, 31.5)	0.49	13.4 (3.36, 53.7)	1.26E-03
	IVW	1.82 (1.20, 2.76)	4.99E-03	2.59 (1.37, 4.92)	3.45E-03	6.79 (2.68, 17.2)	5.21E-05
	Weighted median	2.20 (1.20, 4.04)	0.01	3.87 (1.6, 9.33)	2.61E-03	9.78 (3.24, 29.5)	5.25E-05
	Weighted mode	3.28 (0.92, 11.6)	0.07	6.78 (0.89, 51.7)	0.07	10.4 (3.5, 31.1)	3.14E-04
HPV-positive HNSCC vs controls	MR Egger	6.88 (0.35, 135)	0.21	0.42 (0.02, 9.06)	0.58	4.53 (0.93, 22.0)	0.07
	IVW	2.37 (1.33, 4.24)	3.46E-03	2.60 (1.20, 5.65)	0.02	3.58 (1.27, 10.1)	0.02
	Weighted median	2.41 (1.04, 5.58)	0.04	1.45 (0.47, 4.41)	0.52	3.17 (0.79, 12.7)	0.1
	Weighted mode	2.63 (0.53, 13.19)	0.24	1.03 (0.08, 13.9)	0.98	2.82 (0.73, 10.8)	0.14

HPV, human papillomavirus; MR, Mendelian randomization; OR, odds ratio; CI, confidence interval, IVW, inverse variance-weighted

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

My concerns have been addressed in this revised submission and I have no further comments.

Reviewer #3 (Remarks to the Author):

I would like to thank the authors for providing further clarification and addressing my comments. I have no further major comments but there are still some minor points that the authors may want to change.

1) In the abstract it is claimed that “higher alcohol consumption associated with a stronger risk of HPV-negative HNSCC”. Although the estimate for HPV-negative HNSCC is indeed higher, the confidence intervals overlap. This means that we do not have statistical evidence that the difference observed between the two estimates in not due to chance. Similarly for smoking in the discussion.

2) For my previous point 6. I am still not sure of what the consequence of using the OncoArray was. Maybe the authors can mention the original number of instrument SNPs of the exposure so it is easier to compare to the numbers of SNPs used in the MR.

3) The point of PRS Vs GRS may still confuse readers. A clearer statement will be of benefit given the current use of the PRS term.

Reviewer #1 (Remarks to the Author):

My concerns have been addressed in this revised submission and I have no further comments.

Reviewer #3 (Remarks to the Author):

I would like to thank the authors for providing further clarification and addressing my comments. I have no further major comments but there are still some minor points that the authors may want to change.

1) In the abstract it is claimed that “higher alcohol consumption associated with a stronger risk of HPV-negative HNSCC”. Although the estimate for HPV-negative HNSCC is indeed higher, the confidence intervals overlap. This means that we do not have statistical evidence that the difference observed between the two estimates in not due to chance. Similarly for smoking in the discussion.

- We thank the reviewer for the suggesting regarding the interpretation of the results. Given the lack of statistical difference between the two estimates, we have removed from the abstract and discussion the writing that may suggests a differences in the effects between the two groups and rather highlight the significance of the effects in both groups independently.
- Removed from abstract: “with lifetime smoking having a similar effect in both cancer types and higher alcohol consumption associated with a stronger risk of HPV-negative HNSCC” from abstract.
- Removed from discussion: “While alcohol use was found to be more strongly associated with the risk of developing HPV-negative HNSCC, we observed an independent causal effect in both HPV-positive and HPV-negative HNSCC”

2) For my previous point 6. I am still not sure of what the consequence of using the OncoArray was. Maybe the authors can mention the original number of instrument SNPs of the exposure so it is easier to compare to the numbers of SNPs used in the MR.

- Thank you for your insightful comment. The OncoArray is a DNA microarray specifically designed for cancer research and is used to detect genetic variations associated with cancer risk. The initial number of SNPs genotyped by the OncoArray was 533,631, which was reduced to 486,987 after applying standard quality control procedures. Moreover, the OncoArray chip includes a genome-wide backbone of approximately 250,000 SNPs that tag most common variants, enabling the imputation of additional variants with high quality beyond those directly captured by the chip. This extensive coverage ensures robust data for Mendelian Randomization (MR) analyses and facilitates comparisons across different datasets. We have included the number of SNPs genotyped on the Oncoarray in the methods section.

3) The point of PRS Vs GRS may still confuse readers. A clearer statement will be of benefit given the current use of the PRS term.

- We thank the reviewer for pointing out this issue. We used the term "polygenic risk scores (PRS)" throughout the manuscript, especially for the factorial MR. Specifically, we first constructed the polygenic risk scores for the exposures as instrumental variables for each study participant. Secondly, we performed a one-sample MR of each phenotype on both HPV-positive and HPV-negative HNSCC.
- We modified in the methods section to improve clarity on this topic “The PRS aggregates the effects of multiple genetic variants to estimate an individual's genetic predisposition to the specific exposure in question. The SNPs used to construct the PRS, along with their effect sizes, were the same as those used in the **univariable MR”**