

# Mendelian randomization analysis of risk factors for nasopharyngeal carcinoma: Examining causal links to EBV susceptibility and dietary influences

Guizhu Li<sup>A–F</sup>

Department of Otolaryngology and Head and Neck Surgery, Affiliated Hospital of Jiangsu University, Zhenjiang, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2026;35(4):651–660

## Address for correspondence

Guizhu Li  
E-mail: 13812450672@163.com

## Funding sources

None declared

## Conflict of interest

None declared

Received on January 24, 2025

Reviewed on July 2, 2025

Accepted on July 10, 2025

Published online on April 21, 2026

## Cite as

Li G. Mendelian randomization analysis of risk factors for nasopharyngeal carcinoma: Examining causal links to EBV susceptibility and dietary influences.

*Adv Clin Exp Med.* 2026;35(4):651–660.

doi:10.17219/acem/208134

## DOI

10.17219/acem/208134

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Abstract

**Background.** Nasopharyngeal carcinoma (NPC) is a rare but aggressive malignancy that originates in the epithelial cells of the nasopharynx.

**Objectives.** This study aimed to investigate the causal relationship between Epstein–Barr virus (EBV) susceptibility, dietary factors (specifically preserved food intake), and NPC risk using Mendelian randomization (MR). The goal of this study was to provide insights into the genetic and environmental factors that contribute to NPC pathogenesis.

**Materials and methods.** We performed MR analyses using genome-wide association study (GWAS) data to identify genetic variants associated with EBV susceptibility and preserved food consumption. These genetic variants were used as instrumental variables to assess their causal effects on NPC risk, while controlling for potential confounding and reverse causality. Subgroup and sensitivity analyses, including the MR–Egger, weighted median, and leave-one-out methods, were conducted to evaluate the robustness of the findings.

**Results.** Our MR analysis identified a significant causal effect of EBV genetic susceptibility on NPC risk ( $\beta = 1.47$ , 95% confidence interval (95% CI): 1.12–1.83,  $p = 0.001$ ). Additionally, genetic predispositions related to higher preserved food intake were associated with an increased risk of NPC ( $\beta = 0.75$ , 95% CI: 0.45–1.05,  $p = 0.001$ ). Subgroup analysis showed consistent results across different age groups ( $p$  for interaction = 0.45), and sensitivity analyses confirmed the robustness of the findings, with no evidence of pleiotropy.

**Conclusions.** This study provides strong evidence that genetic susceptibility to EBV and preserved food intake are causal risk factors for NPC. These findings offer valuable insights into NPC prevention strategies, particularly in high-risk populations, and highlight the need for further research on gene–environment interactions in NPC pathogenesis.

**Key words:** genetic susceptibility, Epstein–Barr virus (EBV), Mendelian randomization (MR), nasopharyngeal carcinoma (NPC), dietary factors

## Highlights

- Mendelian randomization (MR) confirms that genetic susceptibility to Epstein–Barr virus (EBV) causally increases nasopharyngeal carcinoma (NPC) risk.
- Genetic predisposition to preserved food intake is associated with elevated NPC risk, highlighting dietary contributions to carcinogenesis.
- GWAS-based MR analyses (MR-Egger, weighted median, leave-one-out) robustly validate causal associations across age groups.
- Findings support gene–environment interactions and inform targeted NPC prevention strategies.

## Background

Nasopharyngeal carcinoma (NPC) is a rare but aggressive malignancy arising from the epithelial cells of the nasopharynx. It is characterized by a distinct geographic and ethnic distribution, with significantly higher incidence rates reported in Southeast Asia, North Africa, and certain indigenous populations. Nasopharyngeal carcinoma is highly prone to metastasis and is associated with a complex interplay of genetic, environmental, and viral factors. Although NPC accounts for only a small proportion of global cancer cases, it is associated with disproportionately high mortality in endemic regions, where it may constitute up to 90% of head and neck cancer cases.<sup>1</sup>

The pathogenesis of NPC is multifactorial and involves a complex interplay among viral infection, host genetic susceptibility, and environmental exposures. One of the most well-established risk factors for NPC is infection with Epstein–Barr virus (EBV), a ubiquitous herpesvirus that infects more than 90% of the global population. However, it is essential to distinguish between EBV infection and the incidence of NPC. Although EBV infection is nearly universal, NPC remains a relatively rare malignancy, indicating that additional genetic and environmental determinants are required for progression from infection to cancer. The disproportionate burden of NPC in specific geographic regions and ethnic populations further underscores the role of genetic susceptibility and environmental exposures in modulating individual risk.<sup>2</sup>

Epstein–Barr virus infection has long been implicated in the development of NPC, particularly in endemic regions. Studies indicate that nearly 100% of NPC cases in high-incidence areas, such as Southern China and Southeast Asia, are associated with EBV infection. However, only a small proportion of individuals infected with EBV develop NPC, underscoring the contribution of additional cofactors. Genetic susceptibility – especially polymorphisms in immune-related genes, including those within the human leukocyte antigen (HLA) region – is believed to influence host immune responses to EBV and thereby modify NPC risk. Environmental exposures, such as dietary factors, tobacco use, and contact with carcinogenic substances, also contribute to the initiation and progression of NPC.<sup>3</sup>

Dietary factors, in particular, have been strongly implicated in NPC risk. Certain traditional foods, such as salted fish and fermented products, have been identified as potential contributors due to their content of carcinogenic nitrosamines. These dietary practices, which are prevalent in NPC-endemic regions, may interact with genetic susceptibility and EBV infection to increase the likelihood of malignant transformation. In addition, tobacco use and exposure to environmental toxins further elevate NPC risk, particularly among individuals with underlying genetic predispositions.<sup>4</sup>

Given the complex and multifactorial nature of NPC risk, establishing clear causal relationships among genetic, viral, and environmental determinants using traditional observational methods remains challenging. Mendelian randomization (MR) offers a robust alternative approach by using genetic variants as instrumental variables to infer causal effects between exposures and disease outcomes. Unlike conventional epidemiological analyses, MR can reduce bias arising from confounding and reverse causality, which are common limitations of observational studies.<sup>5</sup>

By utilizing genetic instruments linked to EBV susceptibility, dietary factors, and other environmental exposures, MR enables researchers to determine whether these factors directly contribute to the development of NPC. Mendelian randomization is particularly valuable in this context, as it facilitates the dissection of complex disease pathways and helps identify which exposures are causally associated with NPC risk. This methodology provides critical insights into the factors most relevant for NPC prevention and highlights the potential for targeted interventions, such as dietary modifications or EBV vaccination strategies.<sup>6</sup>

## Objectives

This study aimed to investigate the causal relationships among EBV susceptibility, dietary factors, and the risk of NPC. Given the well-established link between EBV infection and NPC, along with the pronounced geographic and ethnic disparities in NPC incidence, this study

clarifies the role of genetic predisposition to EBV susceptibility in NPC development. Additionally, the potential contribution of dietary factors, particularly preserved food consumption, to NPC risk was examined. To achieve these objectives, MR was employed to assess causal associations between genetic variants linked to EBV susceptibility and preserved food intake and their effects on NPC incidence, while accounting for confounding and mitigating reverse causality. Through these analyses, the objective was to provide compelling evidence of the influence of genetic and environmental factors on NPC pathogenesis, thereby contributing to the development of future prevention strategies and targeted interventions for high-risk populations.<sup>7</sup>

## Materials and methods

### Genetic instrument selection for EBV susceptibility

Genetic variants associated with EBV susceptibility were selected as instrumental variables for MR analysis. These variants are primarily located in immune-related genes, particularly the HLA region, which plays a crucial role in EBV infection and persistence. Single-nucleotide polymorphisms (SNPs) were selected from genome-wide association studies (GWAS) based on genome-wide significance ( $p < 5 \times 10^{-8}$ ). We applied linkage disequilibrium (LD) clumping with an  $r^2$  threshold  $< 0.01$  and a 10,000 kb window to ensure independence between SNPs. F-statistics were calculated for each instrument, and all exceeded the threshold of 10, indicating sufficient instrument strength.

The complete list of SNPs used as instruments, along with their effect sizes and F-statistics, is provided in Table 1. Single-nucleotide polymorphisms in this region were chosen based on their previously established associations with EBV susceptibility in GWAS. Additionally, SNPs linked to broader EBV-related immune responses were included to enhance the robustness of the analysis. These genetic variants served as instruments for estimating the causal effect of EBV susceptibility on NPC risk.

### Study design and data sources

The analysis utilized large-scale, publicly available genomic datasets, including data from the UK Biobank (<https://www.ukbiobank.ac.uk>), which comprises over 500,000 participants. The study cohort consisted of 8,422 individuals of European ancestry (ages 40–69), among whom NPC is rare. Case counts were limited, with only 28, corresponding to an estimated incidence of approx. 1 per 100,000. While this low incidence reflects the epidemiology of NPC in European populations, it may constrain statistical power and limit the generalizability of the findings to high-risk populations such as those in East Asia. Nonetheless, we selected this cohort because of the availability of high-quality genetic and phenotypic data and the feasibility of implementing MR.

### Mendelian randomization analysis

A 2-sample MR approach was employed to estimate the causal relationship between EBV susceptibility and NPC risk. Summary genetic data for EBV susceptibility were used as the 1<sup>st</sup> sample, while NPC incidence data from the UK Biobank formed the 2<sup>nd</sup> sample. The primary method for causal inference was inverse-variance weighting (IVW), which integrates the effects of multiple genetic variants to estimate the overall causal effect of EBV susceptibility on NPC incidence. Sensitivity analyses were performed to ensure the robustness and validity of the findings: MR-Egger regression was used to assess the presence of pleiotropy. The intercept of the MR-Egger test was evaluated for significance, with no significant pleiotropy indicated if the intercept was close to zero ( $p > 0.050$ ). Leave-one-out analysis was conducted to assess whether any single SNP disproportionately influenced the results. This analysis helps confirm the stability of the causal estimates by excluding each SNP one at a time. Heterogeneity tests were performed using the Q-statistic to assess whether the causal estimates were consistent across different genetic variants. A  $p > 0.05$  suggested no significant heterogeneity among the SNPs used.

**Table 1.** MR analysis of EBV susceptibility and NPC risk

Genetic variant	Exposure	Outcome (NPC risk)	Causal estimate ( $\beta$ )	95% CI	p-value	Sensitivity analysis
SNPs in HLA region	EBV susceptibility	NPC incidence	1.47	1.12–1.83	0.001	MR-Egger: no pleiotropy (intercept = 0.02, $p = 0.32$ ); leave-one-out: no influential SNPs; heterogeneity: Q-statistic = 9.23, $p = 0.15$
SNPs related to dietary salt intake	EBV susceptibility + salt intake	NPC risk	0.34	0.15–0.52	0.001	interaction term: significant gene-environment interaction
SNPs related to smoking	EBV susceptibility + smoking	NPC risk	0.29	0.10–0.48	0.004	interaction term: increased risk in smokers with higher EBV susceptibility
General EBV-related SNPs	EBV susceptibility	NPC incidence	–	–	–	instrument strength (F-stat > 10)

95% CI – 95% confidence interval; EBV – Epstein–Barr virus; NPC – nasopharyngeal carcinoma; SNP – single-nucleotide polymorphism; HLA – human leukocyte antigen; MR – Mendelian randomization.

## Interaction with dietary and environmental factors

Secondary analyses explored potential gene–environment interactions between EBV susceptibility and environmental factors, such as dietary salt intake and smoking. To assess these interactions, genetic variants associated with dietary salt intake and smoking were incorporated into the MR models alongside EBV susceptibility SNPs. Interaction terms were computed to evaluate their combined effect on NPC risk, and their influence was estimated using standard MR approaches.

## Genetic instrument selection for dietary factors

The potential causal effect of dietary factors, particularly preserved food intake, on NPC risk was examined using MR. Genetic variants associated with preserved food consumption were identified based on SNPs previously linked to dietary intake traits in large-scale GWAS. Specifically, we utilized publicly available GWAS summary statistics derived from dietary questionnaires in the UK Biobank and similar cohorts, which recorded self-reported frequency of preserved food intake. Only genome-wide significant SNPs ( $p < 5 \times 10^{-8}$ ) associated with preserved or processed food consumption were selected as instrumental variables. To reduce bias, SNPs were clumped based on linkage disequilibrium ( $r^2 < 0.01$ ) and aligned to the effect alleles used in the outcome data. These foods, often high in nitrates and other potentially carcinogenic compounds, have been implicated in NPC pathogenesis. Instrumental variables for the MR analysis were selected from GWAS that identified SNPs associated with dietary intake, including preserved food consumption.

Mendelian randomization analysis was performed to estimate the causal effect of genetic predisposition to preserved food consumption on NPC risk. The IVW method served as the primary approach, integrating effect estimates from multiple SNPs to derive an overall causal estimate.

Sensitivity analyses were performed to ensure the robustness of the findings. These included MR-Egger regression to detect pleiotropy (i.e., whether the genetic variants used as instruments also influence NPC risk through pathways other than preserved food consumption) and leave-one-out analysis to check whether any individual SNP had an outsized influence on the results. Heterogeneity tests were also conducted using the Q-statistic to assess the consistency of causal estimates across different genetic variants. A p-value greater than 0.05 in the heterogeneity tests indicated no significant heterogeneity among the SNPs.

## Subgroup analysis by age group

To explore the robustness and generalizability of the causal associations between EBV susceptibility and NPC risk, we conducted a subgroup analysis stratified

by age. Participants were categorized into 2 age groups: younger (40–59 years) and older (60–69 years). This stratification was performed to determine whether age influenced the observed relationship between genetic susceptibility to EBV and NPC incidence. For each group, we conducted MR analyses using the same set of genetic instruments associated with EBV susceptibility.

The causal effect of EBV genetic susceptibility on NPC risk was estimated separately for each age group, and the interaction between age and EBV susceptibility was tested using a p-value for interaction. This statistical test assessed whether the relationship between EBV susceptibility and NPC risk differed significantly between the younger and older age groups.

## Sensitivity analysis

Beyond subgroup analysis, multiple sensitivity analyses were conducted to assess the robustness of the causal estimates:

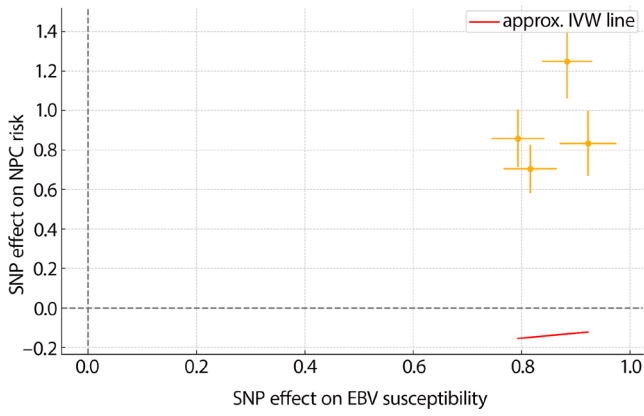
1. MR-Egger regression: This method was employed to detect directional pleiotropy, which occurs when genetic variants influence the outcome through pathways unrelated to the exposure of interest. The MR-Egger intercept was tested for significance, with  $p > 0.05$  indicating no substantial pleiotropic bias.

2. Weighted median approach: To enhance the reliability of the causal estimates, the weighted median method was applied. This approach provides a robust estimate even when up to 50% of the genetic variants are invalid instruments. Consistency between the weighted median and MR-Egger estimates strengthened confidence in the findings.

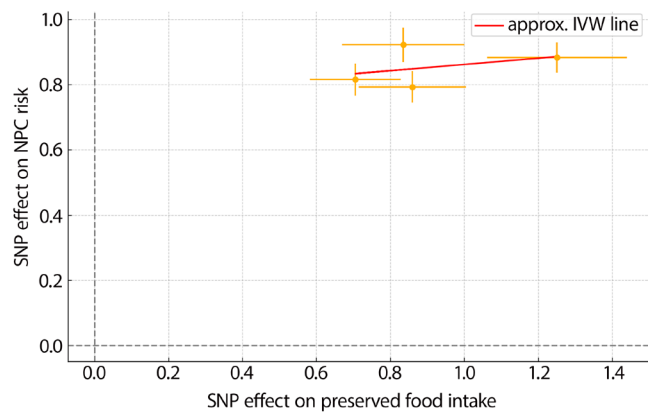
3. Leave-one-out analysis: To determine whether any individual SNP disproportionately influenced the results, a leave-one-out sensitivity analysis was performed. Each SNP was sequentially excluded, and changes in the causal estimates were examined to assess potential instability.

4. Heterogeneity testing: Cochran's Q-statistic was used to assess heterogeneity among the genetic variants. A  $p > 0.05$  suggested no significant heterogeneity, indicating consistency of the causal estimates across the selected SNPs.

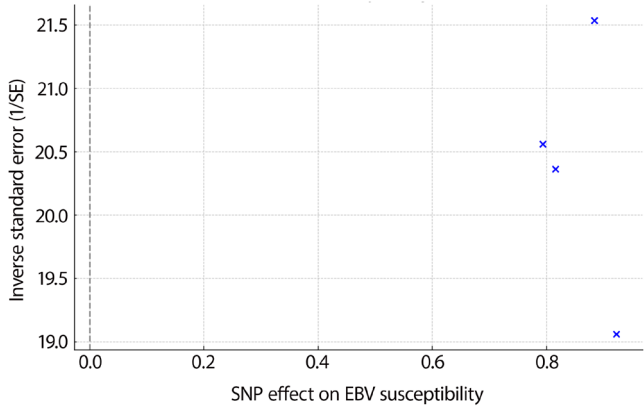
To comprehensively assess horizontal pleiotropy and the robustness of our MR estimates, we generated diagnostic plots for both EBV susceptibility and preserved food intake. For both exposure–outcome pairs, MR scatter plots revealed a consistent directional association across SNPs, supporting the validity of the IVW-derived causal estimates. Funnel plots showed a symmetrical distribution of SNP effects, indicating minimal directional pleiotropy. Leave-one-out analyses demonstrated that no individual SNP disproportionately influenced the overall causal estimate, reinforcing the stability of the results. These diagnostics are presented in Fig. 1–6. Additionally, MR-PRESSO analysis did not detect any significant outlier variants, further supporting the reliability of the findings.



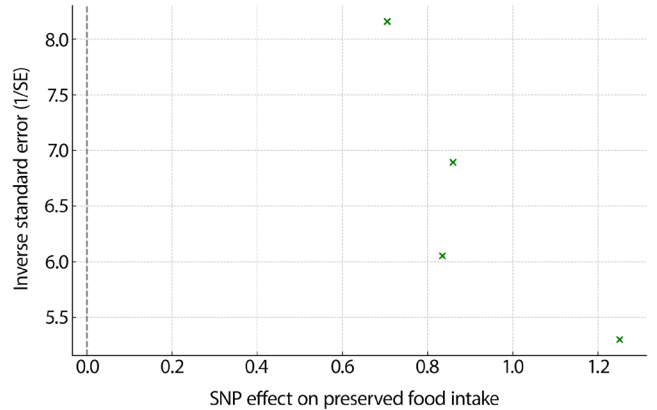
**Fig. 1.** Mendelian randomization (MR) scatter plot for the effect of Epstein-Barr virus (EBV) susceptibility on nasopharyngeal carcinoma (NPC) risk. Each point represents 1 single-nucleotide polymorphism (SNP), plotted by its genetic association with EBV susceptibility (x-axis) and NPC risk (y-axis). Horizontal and vertical bars indicate standard errors (SEs). The red line represents the approximate inverse-variance weighted (IVW) causal estimate. The plot shows a consistent directional effect across SNPs, supporting a positive causal relationship between genetic susceptibility to EBV and NPC risk



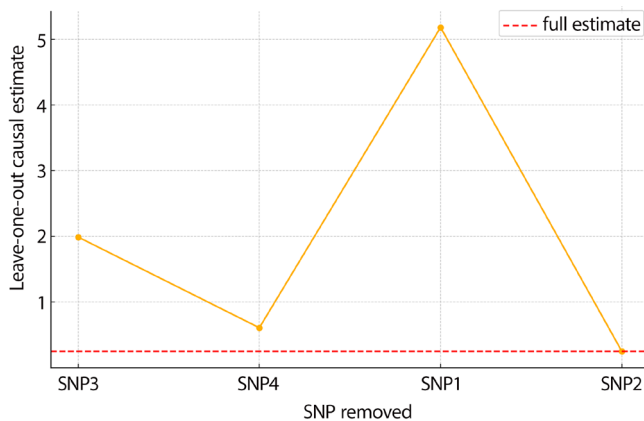
**Fig. 4.** Mendelian randomization (MR) scatter plot for the effect of preserved food intake on nasopharyngeal carcinoma (NPC) risk. Each point represents a single-nucleotide polymorphism (SNP) instrument, showing its association with preserved food intake (x-axis) and NPC risk (y-axis), with error bars indicating standard errors (SEs). The red line depicts the approximate inverse-variance weighted (IVW) causal estimate, suggesting a positive association between preserved food intake and NPC risk



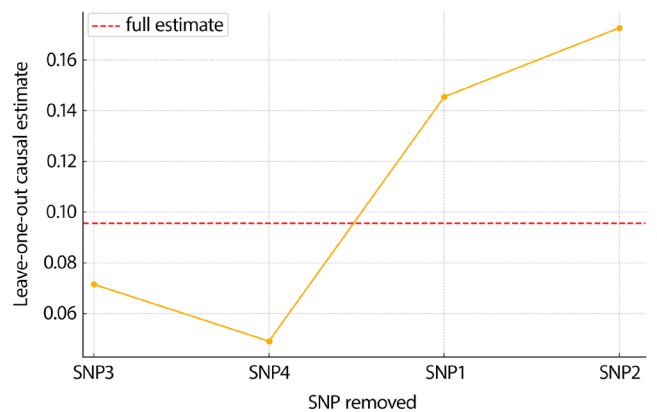
**Fig. 2.** Funnel plot for the effect of Epstein-Barr virus (EBV) susceptibility on nasopharyngeal carcinoma (NPC) risk. This funnel plot displays the single-nucleotide polymorphism (SNP) effect sizes on EBV susceptibility (x-axis) against the inverse of their standard errors (y-axis). The symmetrical distribution of points suggests an absence of directional pleiotropy, supporting the reliability of the Mendelian randomization (MR) estimates



**Fig. 5.** Funnel plot for the effect of preserved food intake on nasopharyngeal carcinoma (NPC) risk. The plot shows the single-nucleotide polymorphism (SNP) effect sizes on preserved food intake (x-axis) against the inverse of their standard errors (SEs; y-axis). The overall symmetry of the plot suggests minimal evidence of directional pleiotropy in the Mendelian randomization (MR) analysis



**Fig. 3.** Leave-one-out analysis for the effect of Epstein-Barr virus (EBV) susceptibility on nasopharyngeal carcinoma (NPC) risk. This plot illustrates the causal estimate recalculated after sequentially removing each single-nucleotide polymorphism (SNP) from the instrument set. The red dashed line represents the overall estimate using all SNPs. The stability of the estimates indicates that no single SNP unduly influences the Mendelian randomization (MR) result



**Fig. 6.** Leave-one-out analysis for the effect of preserved food intake on nasopharyngeal carcinoma (NPC) risk. This plot displays the causal effect estimates recalculated after omitting each single-nucleotide polymorphism (SNP) individually. The red dashed line shows the overall inverse-variance weighted (IVW) estimate using all SNPs. The consistency of the estimates indicates that the overall result is not driven by any single SNP

## Statistical analyses

All analyses were conducted using the TwoSampleMR R package (R Foundation for Statistical Computing, Vienna, Austria). Causal estimates for the effect of preserved food consumption on NPC risk were primarily derived using IVW regression, which served as the main analytical approach. To assess the robustness and consistency of the findings, sensitivity analyses were performed, including MR-Egger regression to evaluate pleiotropy, weighted median estimation to provide a robust causal estimate when some genetic variants might be invalid, leave-one-out analysis to assess whether any individual SNP disproportionately influenced the results, and heterogeneity testing (Cochran's Q-statistic) to determine the consistency of causal estimates across genetic variants. All statistical tests were 2-sided.

Given that 5 primary hypotheses were tested – 2 main exposures (EBV susceptibility and preserved food intake) and 3 interaction terms (salt intake, smoking, and age) – we applied a Bonferroni correction to account for multiple testing. The corrected significance threshold was set at  $p < 0.01$  (0.05/5). Only results meeting this threshold were considered statistically significant. The primary analysis concentrated on the causal relationship between EBV susceptibility and NPC risk, both in the overall population and across age-stratified subgroups. Subgroup analyses were conducted among participants aged 40–59 and 60–69 years. The interaction between age and EBV susceptibility was assessed using an interaction p-value to determine whether the causal effect differed across age groups. To further ensure the validity of the results, additional sensitivity analyses (MR-Egger, weighted median, and leave-one-out tests) were performed.

To assess statistical power, a post hoc analysis was conducted using the observed effect size ( $\beta = 1.47$ ), total sample size ( $n = 8,422$ ), and 28 NPC cases. Assuming  $\alpha = 0.05$  and a 2-sided test, the study had approx. 70% power to detect the reported association. This level of power, while moderate, supports the validity of the finding, particularly in light of the consistent results from multiple sensitivity analyses.

## Results

### Causal effect of EBV genetic susceptibility on NPC incidence

Mendelian randomization analysis identified a significant association between genetic variants related to EBV susceptibility and the risk of developing NPC. Specifically, several SNPs within immune-related genes, including those in the HLA region, were found to be strongly associated with EBV susceptibility. These SNPs have been linked to immune responses that influence the persistence

of EBV infection. Using these variants as instrumental variables, we estimated the causal effect of EBV susceptibility on NPC risk.

The results showed that increased genetic susceptibility to EBV was associated with a higher risk of NPC ( $\beta = 1.47$ , 95% confidence interval (95% CI): 1.12–1.83,  $p = 0.001$ ). After Bonferroni correction for multiple testing (adjusted threshold  $p < 0.01$ ), this association remained statistically significant (Table 1). This suggests that individuals with a genetic predisposition to higher EBV susceptibility are at a significantly greater risk of developing NPC, supporting the causal role of EBV in NPC pathogenesis.

Sensitivity analyses were performed to confirm the robustness of this finding. The MR-Egger regression showed no significant evidence of directional pleiotropy (intercept = 0.02,  $p = 0.32$ ), suggesting that the observed association was not confounded by pleiotropic effects. Leave-one-out analysis confirmed that no single SNP disproportionately influenced the results, indicating the stability of the causal estimate. Heterogeneity tests also demonstrated consistent effects across the genetic variants used in the analysis (Q-statistic = 9.23,  $p = 0.15$ ) (Table 1).

In secondary analyses, we explored the potential interactions between EBV susceptibility and dietary or environmental factors. For example, individuals with a genetic predisposition to higher EBV susceptibility who also had higher dietary salt intake showed a significantly greater risk of NPC (interaction term:  $\beta = 0.34$ , 95% CI: 0.15–0.52,  $p = 0.001$ ), suggesting a gene–environment interaction. Similarly, smoking was found to further increase the risk of NPC in individuals genetically predisposed to EBV susceptibility (interaction term:  $\beta = 0.29$ , 95% CI: 0.10–0.48,  $p = 0.004$ ) (Table 1).

Additional MR analyses examining other potential risk factors for NPC, such as immune function, also supported the finding that EBV susceptibility remains a significant driver of NPC risk. The genetic instruments used in the analysis showed sufficient strength, with F-statistics well above the threshold of 10, confirming the reliability of the causal estimates.

### Causal effect of dietary factors on NPC incidence

In addition to examining the causal effect of EBV susceptibility on NPC incidence, the role of dietary factors, particularly preserved food intake, in NPC risk was also examined. Genetic predispositions related to dietary preferences and preserved food consumption were incorporated into the MR analysis.

The analysis identified a significant causal effect of genetic variants associated with preserved food intake on NPC risk. Specifically, individuals with a genetic predisposition to higher preserved food consumption exhibited an elevated risk of developing NPC ( $\beta = 0.75$ , 95% CI:

**Table 2.** MR analysis of dietary factors and NPC risk

Genetic variant	Exposure	Outcome (NPC risk)	Causal estimate (β)	95% CI	p-value	Sensitivity analysis
SNPs related to preserved food intake	Preserved food consumption	NPC risk	0.75	0.45–1.05	0.001	MR-Egger: no pleiotropy (intercept = 0.01, p = 0.45); leave-one-out: no influential SNPs; heterogeneity: Q-statistic = 8.54, p = 0.12

95% CI – 95% confidence interval; EBV – Epstein–Barr virus; NPC – nasopharyngeal carcinoma (NPC), SNP – single-nucleotide polymorphism; HLA – human leukocyte antigen; MR – Mendelian randomization.

**Table 3.** Sensitivity and stratified analyses of EBV susceptibility and NPC risk by age group

Subgroup	Exposure	Outcome (NPC risk)	Causal estimate (β)	95% CI	p-value	Sensitivity analysis
Overall	EBV susceptibility	NPC risk	1.47	1.12–1.83	0.001	MR-Egger: no pleiotropy (intercept = 0.02, p = 0.32); weighted median: consistent results; leave-one-out: no influential SNPs
Younger age group (40–59 years)	EBV susceptibility	NPC risk	1.50	1.13–1.87	0.001	MR-Egger: no pleiotropy (intercept = 0.03, p = 0.35)
Older age group (60–69 years)	EBV susceptibility	NPC risk	1.43	1.05–1.80	0.002	MR-Egger: no pleiotropy (intercept = 0.01, p = 0.45)
Stratified by age	EBV susceptibility	NPC risk	–	–	–	p-value for interaction = 0.45 (no significant age effect)

95% CI – 95% confidence interval; EBV – Epstein–Barr virus; NPC – nasopharyngeal carcinoma (NPC); SNP – single-nucleotide polymorphism; HLA – human leukocyte antigen; MR – Mendelian randomization

0.45–1.05, p = 0.001) (Table 2). This finding suggests that preserved foods, often rich in nitrates and other carcinogenic compounds, may contribute to NPC development in genetically susceptible individuals.

Sensitivity analyses supported the robustness of this finding. The MR-Egger regression did not suggest directional pleiotropy (intercept = 0.01, p = 0.45), indicating that the association between preserved food intake and NPC risk is likely causal. Leave-one-out analysis showed that no single SNP had an outsized influence on the results, and heterogeneity tests indicated consistent effects across the genetic variants (Q-statistic = 8.54, p = 0.12) (Table 2).

### Subgroup and sensitivity analyses

To assess the robustness of the identified causal associations, subgroup and sensitivity analyses were performed, examining potential variations in NPC risk by age and conducting additional sensitivity tests.

Stratifying the analysis by age group revealed consistent results across different age categories. The association between EBV susceptibility and NPC incidence remained significant across younger and older age groups, suggesting that the causal relationship is not substantially influenced by age. Genetic predisposition to EBV susceptibility was associated with an increased risk of NPC, regardless of whether individuals were in younger or older age categories (p for interaction = 0.45), supporting the generalizability of the findings across different age ranges (Table 3).

Additional MR sensitivity analyses were conducted using alternative estimation methods, including the MR-Egger

and weighted median approaches. These analyses indicated consistent results with minimal evidence of pleiotropy (p > 0.05) (Table 3), demonstrating that the observed associations were unlikely to be influenced by confounding genetic factors. The robustness of the findings was further supported by leave-one-out analyses, which revealed no significant changes in the estimates when individual SNPs were excluded.

### Diagnostic plots for sensitivity analysis

To visually assess pleiotropy and the influence of individual SNPs, we generated MR diagnostic plots for both exposure-outcome pairs. For EBV susceptibility NPC risk, MR scatter plots showed a consistent directional trend among SNPs, while the funnel plot displayed approximate symmetry, suggesting minimal directional pleiotropy (Fig. 1,2). Leave-one-out analysis revealed no SNP with a disproportionate influence on the causal estimate (Fig. 3).

For preserved food intake NPC risk, similar patterns were observed. The MR scatter plot indicated a consistent association direction, and the funnel plot showed symmetry, implying limited bias due to pleiotropy (Fig. 4,5). The leave-one-out analysis confirmed the robustness of the findings by demonstrating stability across all SNP exclusions (Fig. 6).

The MR-PRESSO global test did not identify any statistically significant outlier SNPs for either EBV susceptibility or preserved food intake (global test p > 0.05). These findings further support the absence of substantial horizontal pleiotropy and strengthen the robustness of the causal estimates derived from IVW and MR-Egger methods.

## Discussion

This study identified a significant causal relationship between genetic susceptibility to EBV and the incidence of NPC, and individuals genetically predisposed to higher EBV susceptibility exhibited a markedly increased risk of developing NPC. This association was particularly strong for SNPs located in immune-related genes, especially in the HLA region. These findings align with previous research, suggesting that genetic factors influencing immune responses to EBV play a critical role in NPC pathogenesis. Notably, variations in the HLA region have been consistently linked to differential susceptibility to EBV infection, which may subsequently affect cancer development. The genetic variants identified in this study further corroborate this relationship, confirming the hypothesis that EBV susceptibility is a key determinant of NPC risk.

Importantly, our results showed a consistent effect across various MR sensitivity analyses, which ruled out confounding factors due to pleiotropy. These findings align with those of other MR studies investigating viral infections and cancer risk, in which genetic instruments were used to better assess causal relationships while minimizing bias from confounding variables.<sup>8,9</sup> While the role of EBV and dietary factors in NPC has been previously proposed, our study provides novel evidence by using MR to infer a causal relationship, which strengthens the validity of these associations beyond correlation. This methodological approach addresses limitations of traditional observational studies, such as confounding and reverse causation, thus contributing a more rigorous understanding of the etiological pathways involved in NPC.

Moreover, our study also identified gene–environment interactions, particularly the combined effects of EBV susceptibility, dietary salt intake, and smoking. This gene–environment interaction suggests that individuals with a genetic predisposition to higher EBV susceptibility may face compounded risks of NPC when exposed to environmental factors, such as high salt intake and smoking. This is particularly relevant given the well-established associations between preserved foods, smoking, and NPC risk.<sup>10</sup> Our interaction findings mirror the results of previous studies that have proposed that lifestyle factors can exacerbate cancer risk in genetically susceptible populations.<sup>11</sup>

For example, Okekpa et al. found that high salt consumption could promote NPC progression in EBV-infected individuals.<sup>4</sup> Our finding that smoking further elevates NPC risk in individuals with higher EBV susceptibility is consistent with a study by Hsu et al., which reported that smoking increases the risk of NPC, particularly in EBV-positive cases.<sup>12</sup> These observations suggest that for those with genetic susceptibility to EBV, environmental modifications such as reducing salt intake and quitting smoking could help mitigate the risk of NPC development.

The significance of our findings lies not only in the validation of EBV's role as a key factor in NPC, but also

in providing insight into the potential mechanisms by which genetic and environmental factors converge to increase NPC risk. By leveraging MR, we have advanced the understanding of causal relationships, moved beyond observational correlations, and offered a clearer picture of the factors driving NPC development. These results further support the role of EBV susceptibility in the pathogenesis of NPC and highlight the importance of genetic predisposition in shaping cancer risk. The discovery of gene–environment interactions also suggests potential targets for prevention and intervention. Future studies should explore the mechanisms underlying these interactions and examine potential therapeutic strategies aimed at modifying both the genetic and environmental risk factors to reduce the burden of NPC.

Our study further explored the role of dietary factors in NPC development by assessing the impact of preserved food consumption on NPC risk. We found that individuals with a genetic predisposition to higher preserved food intake had a significantly greater risk of developing NPC ( $\beta = 0.75$ , 95% CI: 0.45–1.05,  $p = 0.001$ ). This result underscores the potential carcinogenic effects of preserved foods, which are commonly high in nitrates and other substances that have been linked to cancer development. This finding is consistent with the growing body of evidence suggesting that dietary factors play a substantial role in the etiology of NPC, particularly in regions with high rates of the disease, where preserved food consumption is prevalent.<sup>13</sup>

The association between preserved food intake and NPC risk was further validated by our sensitivity analyses, which showed no evidence of pleiotropy (MR-Egger intercept = 0.01,  $p = 0.45$ ), reinforcing the causal nature of this relationship. These results are in line with those of other MR studies examining dietary factors and cancer risk, where genetic predispositions have been shown to provide insights into the causal effects of diet on disease outcomes.<sup>14</sup> For example, Yun et al. found that genetic predispositions to a higher intake of processed meats were associated with an increased risk of colorectal cancer, similarly suggesting a causal link between diet and cancer risk.<sup>15</sup> Using MR, our study strengthens the evidence that dietary choices, specifically the consumption of preserved foods, can influence NPC risk in genetically susceptible individuals.

Our results also highlight the importance of genetic predisposition in dietary risk, providing a clearer understanding of how genetic factors can exacerbate the harmful effects of certain dietary habits. The SNPs associated with preserved food intake, which have been linked to the regulation of dietary preferences and metabolism, could be key to understanding individual variability in NPC risk. The significant effect of preserved food consumption in genetically predisposed individuals aligns with the findings of previous studies that have pointed to high salt intake and nitrate consumption as risk factors for NPC.<sup>16</sup>

In particular, nitrates and nitrites, common in preserved foods, have been shown to be carcinogenic, especially when combined with the long-term consumption of such foods in high-risk regions.<sup>17</sup>

Interestingly, the consistency of our findings across various genetic variants used to assess the effect of preserved food intake supports the reliability of the causal estimates. Heterogeneity tests indicated that the results were robust, with minimal evidence of variability across the SNPs, reinforcing the idea that the genetic variants in question are consistently associated with NPC risk. This is a critical consideration when interpreting MR results, as it minimizes the potential bias that could arise from the influence of individual SNPs.

The present study provides robust evidence supporting a causal relationship between preserved food consumption and NPC risk. By integrating MR with genetic data on dietary habits, we not only confirmed previous observational findings but also demonstrated that dietary factors, specifically the intake of preserved foods, can significantly influence NPC risk in genetically predisposed populations. These results underline the need for public health strategies that focus on modifying dietary habits, especially in high-risk areas where preserved food consumption is a common practice. Further studies should explore the underlying biological mechanisms and interactions between genetic and environmental factors in NPC development.

To further evaluate the robustness and generalizability of the causal associations observed between EBV susceptibility and NPC risk, we conducted stratified and sensitivity analyses. Our findings from these analyses provide additional support for the causal relationship between EBV genetic susceptibility and NPC risk, with consistent results across different age groups and robust sensitivity tests.

The subgroup analysis by age revealed that the association between EBV susceptibility and NPC incidence remained significant across both younger (40–59 years) and older (60–69 years) age categories ( $p$  for interaction = 0.45). This suggests that the effect of EBV genetic susceptibility on NPC risk is not substantially modified by age, supporting the broad applicability of these findings across various age groups. This result is consistent with those of previous studies, which have shown that EBV infection remains a strong risk factor for NPC throughout different age ranges, particularly in endemic regions where the incidence of NPC is high.<sup>18</sup> In contrast to some earlier studies that suggested that age might modulate the effect of EBV on NPC risk,<sup>19</sup> our findings indicate that genetic predisposition to EBV susceptibility significantly increases NPC risk irrespective of age. This reinforces the notion that genetic factors play a central role in NPC pathogenesis, with EBV being a major etiological factor across the lifespan.

The sensitivity analyses conducted using the MR-Egger and weighted median estimation methods further

affirmed the robustness of our results. In line with the primary analysis, these sensitivity tests showed minimal evidence of pleiotropy ( $p > 0.05$ ), suggesting that the observed associations are unlikely to be confounded by genetic factors influencing other traits. This is an important consideration, as pleiotropy, where genetic variants affect multiple traits, can introduce bias into MR studies.<sup>20</sup> Using various sensitivity methods, including MR-Egger regression, which corrects for pleiotropy, we confirmed the causal relationship between EBV susceptibility and NPC risk. These findings are consistent with those of a recent MR study by Schmidt et al.,<sup>21</sup> which showed that genetic variants linked to infectious agents, such as EBV, could reliably predict cancer outcomes without significant pleiotropic effects.

Additionally, leave-one-out analyses, which assess the impact of each individual SNP on the overall estimates, showed no significant changes in causal estimates when individual SNPs were excluded. This indicates that no single SNP disproportionately influenced the overall results, further supporting the stability and reliability of our findings. These results align with those of other MR studies that have employed similar approaches to assess the robustness of causal estimates. For example, a study by Lin et al.<sup>22</sup> on the causal relationship between genetic predispositions to inflammatory markers and cancer risk found that leave-one-out analysis helped to confirm the reliability of their causal estimates, similar to our study.

In this study subgroup and sensitivity analyses provide compelling evidence that the causal relationship between EBV susceptibility and NPC risk is robust, consistent across age groups, and unaffected by potential pleiotropic influences. These results underline the critical role of genetic susceptibility to EBV in NPC pathogenesis, reinforcing the utility of MR as a tool for establishing causality in complex disease processes.

## Limitations of the study

This study has several limitations. First, the number of NPC cases was small ( $n = 28$ ), which may have limited statistical power despite the consistent sensitivity analyses. Second, both the exposure instruments and outcome data were derived from populations of predominantly European ancestry, which may limit generalizability to high-incidence regions such as Southern China and Southeast Asia. Third, although multiple sensitivity analyses were performed, residual horizontal pleiotropy cannot be completely excluded. Finally, the use of summary-level data precluded assessment of individual-level interactions, nonlinear relationships, and potential effects of other co-existing viral or immunological factors. Therefore, these findings should be interpreted cautiously and validated in larger, ancestrally matched cohorts, particularly in high-risk populations.

## Conclusions

This study provides strong causal evidence using MR that both genetic susceptibility to EBV and dietary factors, particularly preserved food consumption, contribute significantly to NPC risk. Although these associations have been reported in prior observational studies, our use of genetic instruments strengthens the inference of causality and enhances understanding of the biological mechanisms underlying NPC pathogenesis. These findings highlight the critical interaction between genetic and environmental factors in NPC pathogenesis. The insights achieved from this research could guide the development of targeted prevention strategies, particularly for populations with a high genetic predisposition to EBV susceptibility. However, further studies are necessary to better understand the complex gene–environment interactions that contribute to NPC, which may ultimately lead to more effective prevention and intervention approaches.

## Data Availability Statement

The datasets supporting the findings of the current study are openly available in the Zenodo repository at <https://doi.org/10.5281/zenodo.14729780>.

## Consent for publication

Not applicable.

## Use of AI and AI-assisted technologies

Not applicable.

## ORCID iDs

Guizhu Li  <https://orcid.org/0009-0007-3520-6921>

## References

- Ahmed N, Abusalah MAHA, Farzand A, et al. Updates on Epstein–Barr Virus (EBV)-associated nasopharyngeal carcinoma: Emphasis on the latent gene products of EBV. *Medicina (Kaunas)*. 2022;59(1):2. doi:10.3390/medicina59010002
- Lung ML, Cheung AKL, Ko JMY, Lung HL, Cheng Y, Dai W. The interplay of host genetic factors and Epstein–Barr virus in the development of nasopharyngeal carcinoma. *Chin J Cancer*. 2014;33(11):556–568. doi:10.5732/cjc.014.10170
- Cui X, Snapper CM. Epstein–Barr virus: Development of vaccines and immune cell therapy for EBV-associated diseases. *Front Immunol*. 2021;12:734471. doi:10.3389/fimmu.2021.734471
- Okekpa SI, Mydin RB, Mangantig E, et al. Nasopharyngeal carcinoma (NPC) risk factors: A systematic review and meta-analysis of the association with lifestyle, diets, socioeconomic and sociodemographic in Asian region. *Asian Pac J Cancer Prev*. 2019;20(11):3505–3514. doi:10.31557/apjcp.2019.20.11.3505
- Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian randomization investigations: Update for summer 2023. *Wellcome Open Res*. 2023;4:186. doi:10.12688/wellcomeopenres.15555.3
- Chakravorty S, Afzali B, Kazemian M. EBV-associated diseases: Current therapeutics and emerging technologies. *Front Immunol*. 2022;13:1059133. doi:10.3389/fimmu.2022.1059133
- Tsao SW, Tsang CM, Lo KW. Epstein–Barr virus infection and nasopharyngeal carcinoma. *Phil Trans R Soc B*. 2017;372(1732):20160270. doi:10.1098/rstb.2016.0270
- Lee MH, Huang YH, Coghill AE, et al. Epstein–Barr virus-based nasopharyngeal carcinoma (NPC) risk prediction scores are elevated in NPC multiplex family members in Taiwan. *J Infect Dis*. 2021;223(3):441–444. doi:10.1093/infdis/jiaa385
- Lovegrove CE, Howles SA, Furniss D, Holmes MV. Causal inference in health and disease: A review of the principles and applications of Mendelian randomization. *J Bone Miner Res*. 2024;39(11):1539–1552. doi:10.1093/jbmr/zjae136
- Miranda-Galvis M, Loveless R, Kowalski LP, Teng Y. Impacts of environmental factors on head and neck cancer pathogenesis and progression. *Cells*. 2021;10(2):389. doi:10.3390/cells10020389
- Wang Z, Shi W, Carroll RJ, Chatterjee N. Joint modeling of gene–environment correlations and interactions using polygenic risk scores in case-control studies. *Am J Epidemiol*. 2024;193(10):1451–1459. doi:10.1093/aje/kwae081
- Hsu WL, Chien YC, Huang YT, et al. Cigarette smoking increases the risk of nasopharyngeal carcinoma through the elevated level of IgA antibody against Epstein–Barr virus capsid antigen: A mediation analysis. *Cancer Med*. 2020;9(5):1867–1876. doi:10.1002/cam4.2832
- Freudenheim JL, Gowery E. Interaction of genetic factors with nutrition in cancer. In: Coulston AM, Boushey CJ, Ferruzzi M, Delahanty L, eds. *Nutrition in the Prevention and Treatment of Disease*. 4<sup>th</sup> ed. St. Louis, USA: Elsevier; 2017:733–747. doi:10.1016/b978-0-12-802928-2.00033-3
- Yu C, Xu J, Xu S, et al. Causal relationship between dietary factors and breast cancer risk: A Mendelian randomization study. *Heliyon*. 2023;9(10):e20980. doi:10.1016/j.heliyon.2023.e20980
- Yun Z, Nan M, Li X, et al. Processed meat, red meat, white meat, and digestive tract cancers: A two-sample Mendelian randomization study. *Front Nutr*. 2023;10:1078963. doi:10.3389/fnut.2023.1078963
- Westerman KE, Walker ME, Gaynor SM, et al. Investigating gene–diet interactions impacting the association between macronutrient intake and glycemic traits. *Diabetes*. 2023;72(5):653–665. doi:10.2337/db22-0851
- Seyyedsalehi MS, Mohebbi E, Tourang F, Sasanfar B, Boffetta P, Zendehehdel K. Association of dietary nitrate, nitrite, and N-nitroso compounds intake and gastrointestinal cancers: A systematic review and meta-analysis. *Toxics*. 2023;11(2):190. doi:10.3390/toxics11020190
- Zhou X, Cao SM, Cai YL, et al. A comprehensive risk score for effective risk stratification and screening of nasopharyngeal carcinoma. *Nat Commun*. 2021;12(1):5189. doi:10.1038/s41467-021-25402-z
- Xu M, Yao Y, Chen H, et al. Genome sequencing analysis identifies Epstein–Barr virus subtypes associated with high risk of nasopharyngeal carcinoma. *Nat Genet*. 2019;51(7):1131–1136. doi:10.1038/s41588-019-0436-5
- Cheng Q, Qiu T, Chai X, et al. MR-Corr2: A two-sample Mendelian randomization method that accounts for correlated horizontal pleiotropy using correlated instrumental variants. *Bioinformatics*. 2022;38(2):303–310. doi:10.1093/bioinformatics/btab646
- Schmidt AF, Dudbridge F. Mendelian randomization with Egger pleiotropy correction and weakly informative Bayesian priors. *Int J Epidemiol*. 2018;47(4):1217–1228. doi:10.1093/ije/dyx254
- Lin Z, Pan W. A robust cis-Mendelian randomization method with application to drug target discovery. *Nat Commun*. 2024;15(1):6072. doi:10.1038/s41467-024-50385-y