



Regional dietary patterns and risk of primary liver cancer in anhui, china: a population-based case-control study

Yu Chang , Jie Chen, Zhuang Zhang, Wanshui Yang

Keywords:

Primary liver cancer, dietary patterns, case-control study, nutritional epidemiology

Citation: Chang Y, Chen J, Zhang Z, Yang W. Regional dietary patterns and risk of primary liver cancer in anhui, china: a population-based case-control study. *Hepatoma Res.* 2026;12:26. <https://dx.doi.org/10.20517/2394-5079.2026.13>

Received: 31 Jan 2026

First Decision: 18 Mar 2026

Revised: 18 Apr 2026

Accepted: 22 Apr 2026

Published: 8 Jun 2026

Academic Editor:

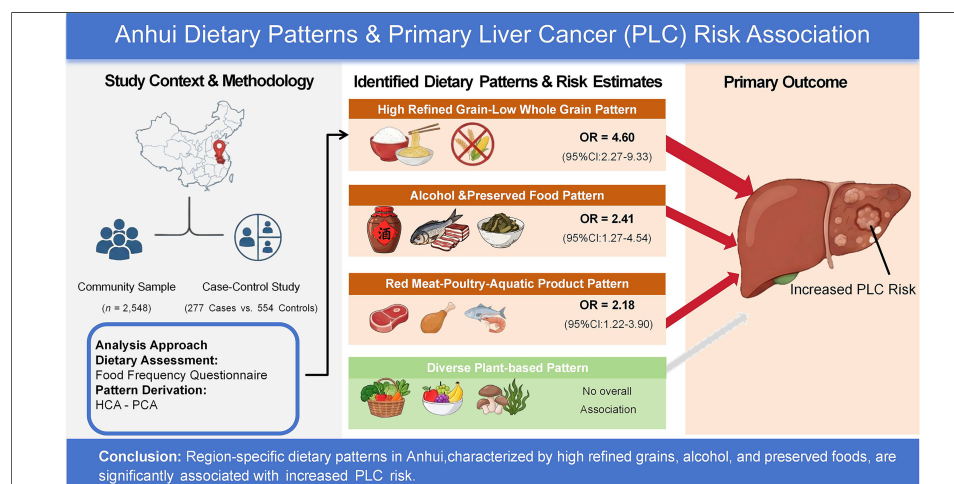
Guangwen Cao

Copy Editor:

Ting-Ting Hu

Production Editor:

Ting-Ting Hu



Abstract

Aim: Primary liver cancer (PLC) remains an important public health concern in Anhui Province, China. Whether dietary habits distinctive to this region are associated with the risk of PLC remains to be investigated. This study aimed to identify characteristic dietary patterns among the adult population in Anhui and examine their associations with PLC risk.

Methods: Using data from a geographically representative community sample ($n = 2,548$), dietary patterns were derived using hierarchical clustering and principal component analysis based on food frequency questionnaires and were structurally validated. A 1:2 matched case-control study (277 PLC cases, 554 controls) was used to assess the associations between these patterns and PLC risk.

Results: Four dietary patterns were identified. The high refined grain-low whole grain pattern [odds ratio (OR) = 4.60; 95% confidence interval (CI): 2.27, 9.33], the alcohol and preserved food co-consumption pattern (OR = 2.41; 95%CI: 1.27, 4.54), and the red meat-poultry-aquatic product pattern (OR = 2.18; 95%CI: 1.22, 3.90) were associated with



Department of Nutrition, Center for Big Data and Population Health of IHM, School of Public Health, Anhui Medical University, Hefei 230032, Anhui, China.

Correspondence to: Dr. Zhuang Zhang, Dr. Wanshui Yang, Department of Nutrition, Center for Big Data and Population Health of IHM, School of Public Health, Anhui Medical University, Hefei 230032, Anhui, China. E-mail: zhzh_ahmu@outlook.com; wanshuiyang@gmail.com

increased PLC risk. In sensitivity analyses, red meat and Chinese liquor appeared to be primary risk contributors within their respective patterns, whereas whole grains, fruits, vegetable oils, and tea appeared to have protective associations. The diverse plant-based pattern was not significantly associated with PLC risk.

Conclusion: Several dietary patterns reflecting local consumption habits in Anhui were associated with PLC risk. These findings suggest the value of region-specific dietary assessment in elucidating lifestyle contributors to PLC.

INTRODUCTION

Primary liver cancer (PLC) is a major public health challenge in Anhui Province, where incidence and mortality rates have consistently exceeded the national average^[1,2]. While hepatitis B virus (HBV) vaccination and control programs have contributed to a national decline in PLC, the persistent regional disparity suggests the potential role of other modifiable factors, with diet being of particular interest^[3].

The dietary landscape in Anhui is shaped by its regional culinary traditions, notably Huizhou cuisine^[4]. Characteristic features include the frequent use of cooking oils, strong seasoning, cooking methods such as braising and stewing that often involve prolonged heating, and regular consumption of preserved foods, including salted meats and pickled vegetables^[5-7]. In addition, the region has a rich tea culture and a strong tradition of soybean product consumption^[5,8]. These dietary patterns may collectively influence long-term health outcomes, yet their association with PLC risk has not been evaluated using a dietary pattern approach.

Although prior studies have examined isolated dietary components, such as red meat, alcohol, or specific nutrients, in relation to PLC^[9-13], these approaches may overlook the synergistic and cumulative effects of foods consumed in combination. While established dietary patterns (e.g., Mediterranean or Western-type diets) have been linked to PLC in other populations, region-specific patterns reflecting local dietary practices in Anhui remain poorly characterised and underexplored in etiological research^[14,15]. Notably, Anhui Province reports a slightly higher PLC burden than the national average^[1,2], yet the underlying reasons for this disparity remain unclear. Given the high and persistent burden of PLC in Anhui, identifying region-specific dietary patterns that may influence PLC risk is warranted.

To address this gap, the present study aimed to identify characteristic “Huizhou-style” dietary patterns among the adult population in Anhui using hierarchical clustering analysis (HCA) and principal component analysis (PCA). This approach was designed to capture sparse and interpretable patterns from locally consumed food groups, drawing conceptually on the rationale of the Treetlet Transform framework^[16,17]. The structural consistency and stability of the derived patterns were further evaluated in an internal validation set generated by randomly splitting the study population in a 3:1 ratio. These empirically derived patterns were subsequently examined in a matched case-control study to assess their associations with PLC risk. By integrating local dietary context into nutritional epidemiology, this work provides insights into potentially modifiable dietary factors related to the elevated PLC burden in Anhui.

METHODS

Study population

Population for dietary pattern derivation and validation

Dietary patterns were derived from a population-based sample of 3,080 adults in Anhui Province, using data collected from July 2020 to July 2022. The ethics review board of Anhui Medical University approved the study protocol (Protocol Number: 20200897). Of these participants, 2,942 were from the Anhui Liver Disease Study, and the remaining 138 were volunteers from various regions in Anhui^[18,19]. Individuals with abnormal

energy intake ($> 3,500$ or < 600 kcal/day for females and $> 4,200$ or < 800 kcal/day for males; $n = 532$) were excluded, leaving 2,548 participants for analysis. The sample was randomly split into a discovery set ($n = 1,911$, 75%) and an internal validation set ($n = 637$, 25%) to assess structural consistency and stability of the identified patterns, as shown in [Supplementary Figure 1](#).

Participant selection for the case-control study

A case-control study was conducted to assess the associations between the derived dietary patterns and PLC risk, as illustrated in [Figure 1](#). PLC case data were obtained from multiple hospitals across the central, western, and southern regions of Anhui Province via the Anhui Liver, Gastric, and Colorectal Cancer Study between October 2020 and July 2024. The ethics review board of Anhui Medical University approved the study protocol (Protocol Number: 81250626). PLC diagnosis was confirmed based on histopathological evidence, computed tomography, or magnetic resonance imaging in accordance with national clinical guidelines. To reduce bias, patients who did not complete the food frequency questionnaire (FFQ, $n = 17$), reported substantial changes in dietary habits within the past year ($n = 263$), or had abnormal energy intake ($n = 27$) were excluded. Consequently, 277 eligible cases were included in the final analysis.

Controls were selected from the aforementioned community population. After excluding 532 individuals with abnormal energy intake, 500 with missing HBV status, and 42 with any history of malignancy, a total of 2,006 potential controls remained. For each case, 2 controls were randomly selected from these individuals and matched on gender and age (± 2 years), yielding a 1:2 matched case-control dataset.

Dietary assessment

Baseline dietary information was collected using a regionally validated 141-item FFQ designed for the Anhui population^[19,20]. The FFQ was administered through face-to-face interviews by trained researchers. Participants reported the average frequency of consumption for a standardized portion size of each food item based on their dietary habits over the past year, with frequencies categorized into 9 levels ranging from “never” to “3 times per day”. Average daily food intake was estimated by multiplying the consumption frequency by the specified portion size for each food item.

Food group

The food items were aggregated into 20 predefined food groups based on nutritional similarity and culinary usage in the Anhui region (e.g., “refined grains” and “Chinese liquor”). Self-reported preference for salty taste was also included as a variable in the pattern analysis. [Supplementary Table 1](#) shows the description of the 20 food groups and salty taste preference.

Covariate assessment

Information on age, gender, education level, smoking status, alcohol consumption, physical activity, and annual household income was collected via a self-report questionnaire. Body mass index (BMI) was calculated as weight (kg) divided by height squared (m^2). Physical activity was quantified as metabolic equivalent of tasks (MET)-hours/week. Diabetes status was defined as a self-reported physician diagnosis of diabetes or a fasting blood glucose level ≥ 7.0 mmol/L. HBV infection status was determined at baseline using an enzyme-linked immunosorbent assay.

Post hoc power analysis

A post hoc power calculation was performed to assess the statistical power of the study. Based on the observed sample size (277 cases and 554 controls), the exposure prevalence in controls for the highest-*vs.*-lowest tertile comparison ($p_0 = 0.50$), and a two-sided α of 0.05, the estimated power ranged from 80.9% to 87.2% for detecting an odds ratio (OR) of 1.7 across the 4 dietary patterns.

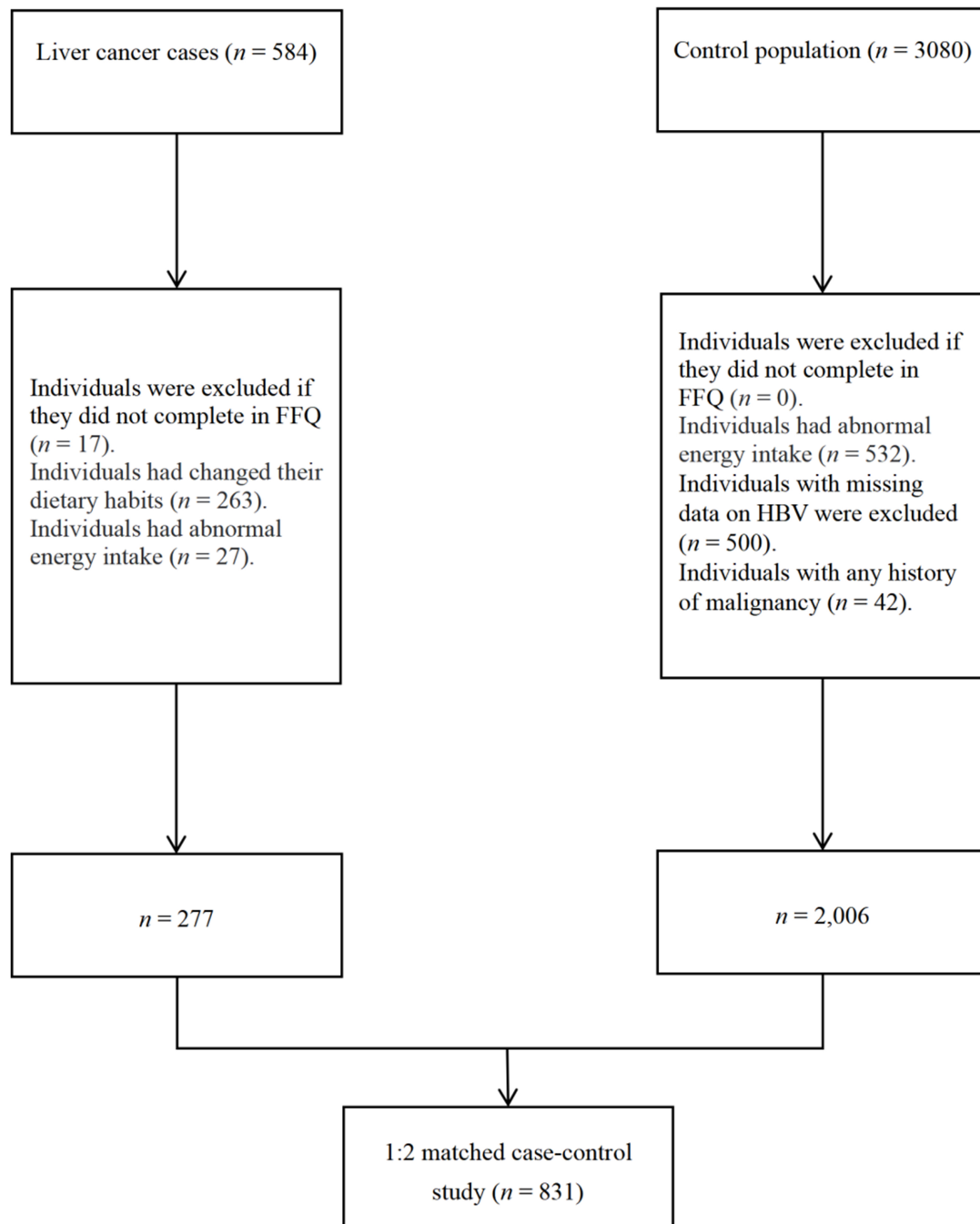


Figure 1. Flowchart of participant selection for the 1:2 matched case-control study. Abnormal energy intake was defined as > 3,500 or < 600 kcal/day for females and > 4,200 or < 800 kcal/day for males. Controls were individually matched to cases by age and gender. FFQ: Food frequency questionnaire; HBV: hepatitis B virus; PLC: primary liver cancer.

Statistical analysis

Dietary pattern derivation and validation

To account for differences in energy intake, food group intakes were adjusted using the nutrient density method (g/1,000 kcal)^[21]. The adjusted variables were then log-transformed to improve normality and reduce the influence of outliers^[22]. Finally, all variables (including food groups and salty taste preference) were standardized to z-scores to ensure equal variance contribution in the subsequent pattern extraction^[23].

Dietary patterns were derived from the discovery set using a hybrid HCA-PCA approach designed to yield sparse and interpretable loading structures, informed by the Treelet Transform framework^[17]. Specifically, the Pearson correlation matrix among dietary variables was computed, with inter-variable distance defined as $1 - |r|$. Variables were then grouped using Ward's hierarchical clustering. For each resulting cluster, suitability for local PCA was assessed using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test; clusters with $KMO \geq 0.5$ and Bartlett's test $P < 0.001$ were subjected to local PCA, and the first principal component (PC1) was retained as the cluster-specific pattern score^[23]. Non-zero loadings were restricted to variables within each cluster to enhance interpretability. To ensure consistent interpretation across resampling procedures, the PC1 loading vector was sign-aligned by constraining the element with the largest absolute loading to be positive. The optimal number of patterns (k) was determined based on average silhouette width, variance explained, and nutritional interpretability, yielding 4 stable dietary patterns^[24].

The consistency and stability of the identified dietary patterns were evaluated through a two-step approach. First, to assess internal consistency, dietary patterns were independently derived in the validation set ($n = 637$) using the same approach applied to the discovery set. The resulting loading matrices were compared with those from the discovery set using Tucker's Congruence Coefficient (TCC), which quantifies the cosine similarity between factor loadings and is suitable for comparing sparse structures. A TCC threshold of ≥ 0.95 was used to define high structural congruence between the 2 independent sub-samples^[25]. Second, to assess internal stability, a resampling analysis was conducted within the discovery set. This involved 200 iterations of random subsampling (drawing 80% of the sample without replacement). Pattern stability was assessed using the mean Jaccard coefficient, with values > 0.60 indicating acceptable stability^[26].

Case-control study analysis

Baseline characteristics were summarised using medians and interquartile ranges (IQRs) for continuous variables and frequencies for categorical variables. Differences between groups were assessed using the Wilcoxon rank-sum test for continuous variables and the chi-square test for categorical variables. Missing covariate data ($< 10\%$ for any variable) were imputed using the multivariate imputation by chained equations, assuming data were missing at random. Five imputed datasets were generated, and regression analyses were performed in each imputed dataset with estimates combined using Rubin's rules. For the case-control analysis, individual food group intakes were first standardised using the mean and standard deviation parameters estimated from the discovery set. For each participant, the score for a given dietary pattern was calculated as the sum of the standardised intakes of the food groups multiplied by their corresponding loading coefficients derived from the discovery set^[27]. Higher scores indicated greater adherence to the corresponding dietary pattern. The association between dietary patterns and PLC risk was evaluated using conditional logistic regression to estimate ORs and 95% confidence intervals (CIs), accounting for the 1:2 matched design. Dietary pattern scores were categorised into tertiles according to the distribution among controls, and linear trends were tested using the median score of each tertile. Model 1 was unadjusted, while Model 2 was adjusted for education, annual household per capita income, energy intake, marital status, smoking status, alcohol drinking, physical activity, BMI, diabetes, and HBV infection at baseline. Additionally, a pattern-level component-exclusion sensitivity analysis was performed to assess the robustness of the associations. Specifically, pattern scores were recalculated by setting the loading coefficient of one food group to zero while retaining the original weights for all other groups^[28]. The multivariable models were then re-estimated using these modified scores to obtain the corresponding ORs. Additional HBV-negative subgroup and interaction analyses were conducted according to HBV infection status using unconditional logistic regression with Firth's penalised likelihood.

Table 1. Loadings of food groups, variance explained by dietary patterns in the discovery set ($n = 1,911$)^{1,2}

Dietary pattern	Variance explained (%)	Description	Loading
Dietary pattern1	12.70	High refined grain-low whole grain pattern	Whole grains (-0.71), refined grains (0.71)
Dietary pattern2	8.50	Diverse plant-based pattern	Tubers (0.37), dark vegetables (0.34), cucurbitaceous vegetables (0.34), allium vegetables (0.24), legumes and legume products (0.40), mushrooms and algae (0.41), fruits (0.34), eggs (0.19), dairy products (0.22), nuts and seeds (0.24)
Dietary pattern3	8.48	Red meat-poultry-aquatic product pattern	Red meat (0.56), poultry (0.58), Fish and shellfish (0.51), vegetable oils (-0.32)
Dietary pattern4	6.89	Alcohol and preserved food co-consumption pattern	Chinese liquor (0.29), salty taste preference (0.02), tea (0.33), pickled vegetables (0.61), salt-cured meat and fish (0.66)
Cumulative explained variance	36.57		

¹Values represent the sparse loading coefficients for each variable within the identified patterns. Positive loadings indicate a direct association with the pattern, while negative loadings indicate an inverse association; ²Salty taste preference was included as a standardised categorical variable in the model.

All statistical analyses were performed using R software (version 4.5.0). Analyses were conducted primarily using the tidyverse, cluster, survival, psych, mice, tableone, ggplot2, and epiR packages. A two-sided $P < 0.05$ was considered statistically significant.

RESULTS

Characteristics of the Anhui population

A total of 2,548 participants were included, comprising 1,911 in the discovery set and 637 in the validation set. The 2 sets were generally comparable in baseline characteristics. The proportion of females was similar in the discovery and validation sets (59.6% vs. 59.2%), and no significant differences were observed in age, education level, annual household per capita income, marital status, smoking status, alcohol drinking, BMI, energy intake, physical activity, or diabetes prevalence. (all $P > 0.05$) [[Supplementary Table 2](#)].

Dietary patterns derived in the discovery set

We identified 4 dietary patterns in the discovery set ($n = 1,911$) [[Table 1](#)]. The clustering solution showed a low average silhouette width (0.07) [[Supplementary Figure 2](#)]. The KMO values for the 4 variable clusters were 0.50, 0.81, 0.66, and 0.52, respectively, and Bartlett's tests were all significant ($P < 0.001$), supporting the suitability of local PCA within each cluster. These 4 dietary patterns explained 12.70%, 8.50%, 8.48%, and 6.89% of the total dietary variability, respectively, and 36.57% of the total variance in food intake overall. The first dietary pattern, labeled the high refined grain-low whole grain pattern, was characterised by high positive loadings for refined grains (0.71) and a negative loading for whole grains (-0.71). The second dietary pattern, designated the diverse plant-based pattern, was characterised by high intakes of dark vegetables (0.34), cucurbitaceous vegetables (0.34), legumes and legume products (0.40), and mushrooms and algae (0.41). The third dietary pattern, termed the red meat-poultry-aquatic product pattern, featured high loadings for poultry (0.58), red meat (0.56), and fish and shellfish (0.51), along with a negative loading for vegetable oils (-0.32). The fourth dietary pattern, the alcohol and preserved food co-consumption pattern, was characterised by the co-consumption of salt-cured meat (0.66), pickled vegetables (0.61), tea (0.33), and Chinese liquor (0.29).

Pattern consistency and internal stability

The consistency and stability of the derived dietary patterns were confirmed [[Supplementary Table 3](#)]. In the validation set ($n = 637$), the patterns showed high consistency compared with those in the discovery set, with

Table 2. Characteristics of PLC cases and matched controls in the Anhui population^{1,2,3}

	Cases	Controls	P	SMD
No. of participants	277	554		
Female (%)	23.5	23.5	1.00	< 0.01
Age (years)	58 (53, 67)	58 (53, 67)	0.91	< 0.01
Education (%)			< 0.01	0.35
Primary school or below	26.4	40.4		
Junior high school	58.8	46.2		
Senior high school or above	13.4	13.4		
NA	1.4	0.0		
Income (%)			< 0.01	0.26
< 10,000 Yuan	44.8	48.0		
10,000-20,000 Yuan	25.6	26.2		
> 20,000 Yuan	23.5	24.5		
NA	6.1	1.3		
Married (%)	91.0	86.3	0.10	0.17
Smoking status (%)			0.65	0.06
Never smokers	49.5	48.6		
Former or current smokers	49.5	50.9		
NA	1.1	0.5		
Alcohol drinking (%)			0.67	0.10
Never or light drinkers	67.5	66.6		
Moderate drinkers	16.6	17.0		
Heavy drinkers	13.4	12.3		
NA	2.5	4.2		
BMI (kg/m²)	22.8 (20.6, 25.4)	24.6 (22.7, 27.1)	< 0.01	0.56
Energy intake (kcal/day)	1,838 (1,420, 2326)	2,601 (2,072, 3,181)	< 0.01	0.93
Physical activity (METs-h/week)	94 (59, 154)	129 (83, 206)	< 0.01	0.35
HBV infection (%)	45.1	5.2	< 0.01	1.03
Diabetes (%)	6.1	14.8	< 0.01	0.29

¹Continuous variables were presented as median (IQR), while categorical variables are presented as %; ²Values of polytomous variables may not sum to 100% due to missing values or rounding; ³P values were derived from the Wilcoxon rank-sum test for continuous variables and the chi-square test for categorical variables. BMI: Body mass index; HBV: hepatitis B virus; MET: metabolic equivalent of task; IQR: interquartile range; PLC: primary liver cancer.

TCC values of 1.00, 0.98, 1.00, and 0.99, respectively. Furthermore, internal stability assessed via iterative subsampling in the discovery set yielded Jaccard coefficients ranging from 0.62 to 1.00, indicating acceptable to high stability.

Characteristics of cases and controls

Following age and gender matching, the final analysis included 277 PLC cases and 554 controls. The 2 groups had identical gender distributions (23.5% female) and comparable median ages (58 years; IQR: 53, 67).

Compared with controls, cases had higher education levels and lower household income, BMI (22.8 kg/m² vs. 24.6 kg/m²), total energy intake (1,838 kcal/day vs. 2,601 kcal/day), and physical activity (94 MET-hours/week vs. 129 MET-hours/week). Regarding clinical history, the prevalence of HBV infection was

Table 3. Associations between dietary pattern scores and the risk of PLC in the 1:2 matched case-control study (n = 831)^{1,2,3}

	OR (95%CI)			<i>P</i> _{trend}
	T1	T2	T3	
High refined grain-low whole grain pattern				
Model 1	1 (reference)	3.44 (2.13, 5.55)	4.83 (3.02, 7.72)	< 0.01
Model 2	1 (reference)	3.41 (1.63, 7.13)	4.60 (2.27, 9.33)	< 0.01
Diverse plant-based pattern				
Model 1	1 (reference)	0.87 (0.60, 1.28)	1.69 (1.19, 2.40)	< 0.01
Model 2	1 (reference)	0.66 (0.30, 1.43)	1.73 (0.86, 3.46)	0.12
Red meat-poultry-aquatic product pattern				
Model 1	1 (reference)	1.38 (0.95, 2.02)	2.12 (1.45, 3.11)	< 0.01
Model 2	1 (reference)	1.41 (0.78, 2.56)	2.18 (1.22, 3.90)	< 0.01
Alcohol and preserved food co-consumption pattern				
Model 1	1 (reference)	0.66 (0.44, 0.98)	1.64 (1.14, 2.35)	< 0.01
Model 2	1 (reference)	0.79 (0.41, 1.50)	2.41 (1.27, 4.54)	< 0.01

¹Model 1 was unadjusted. Model 2 was adjusted for education (primary school or below, junior high school, senior high school or above), annual household per capita income (< 10,000 Yuan, 10,000-20,000 Yuan, > 20,000 Yuan), energy intake (kcal/day, continuous), marital status, smoking status (never, former or current), alcohol drinking (never or light, moderate, heavy), physical activity (continuous), BMI (continuous), diabetes, HBV infection at baseline; ²For each participant, the score for a given dietary pattern was calculated as the sum of the standardised intakes of the food groups multiplied by their corresponding loading coefficients derived from the discovery set. Higher scores indicated greater conformity to the corresponding dietary pattern; ³The trend tests were conducted using the median value of each tertile as a continuous variable. BMI: Body mass index; OR: odds ratio; CI: confidence interval; PLC: primary liver cancer.

higher in cases than in controls (45.1% vs. 5.2%), whereas the prevalence of diabetes was lower in cases than in controls (6.1% vs. 14.8%) (all $P < 0.05$). No significant differences were observed in marital status, smoking status, or alcohol consumption between the 2 groups (all $P > 0.05$) [Table 2].

Association between dietary patterns and PLC risk

The associations between the 4 identified dietary patterns and PLC risk are presented in Table 3. In the multivariable-adjusted model (Model 2), comparing the highest vs. lowest tertiles, 3 of the 4 dietary patterns showed positive associations with PLC risk: the high refined grain-low whole grain pattern (OR = 4.60; 95%CI: 2.27, 9.33), the alcohol and preserved food co-consumption pattern (OR = 2.41; 95%CI: 1.27, 4.54), and the red meat-poultry-aquatic product pattern (OR = 2.18; 95%CI: 1.22, 3.90) (all $P_{trend} < 0.05$). Conversely, the diverse plant-based pattern showed no statistical association (OR = 1.73; 95%CI: 0.86, 3.46). A Bayesian information criterion (BIC)-based Bayes factor analysis for this pattern yielded $BF_{01} = 5.68$, indicating that the data provided greater support for the null model.

Component-exclusion sensitivity analyses

Component exclusion sensitivity analyses [Figure 2] showed that the observed associations were largely driven by specific food groups. For the high refined grain-low whole grain pattern, excluding refined grains attenuated the association, whereas excluding whole grains strengthened it. For the diverse plant-based pattern, a positive association was observed after excluding legumes and legume products, mushrooms and algae, eggs, or fruits. The association for the red meat-poultry-aquatic product pattern was no longer observed after excluding vegetable oils, but was strengthened when poultry was excluded. Similarly, for the alcohol and preserved food co-consumption pattern, the association was attenuated after excluding Chinese liquor or pickled vegetables, but was strengthened when tea was excluded.

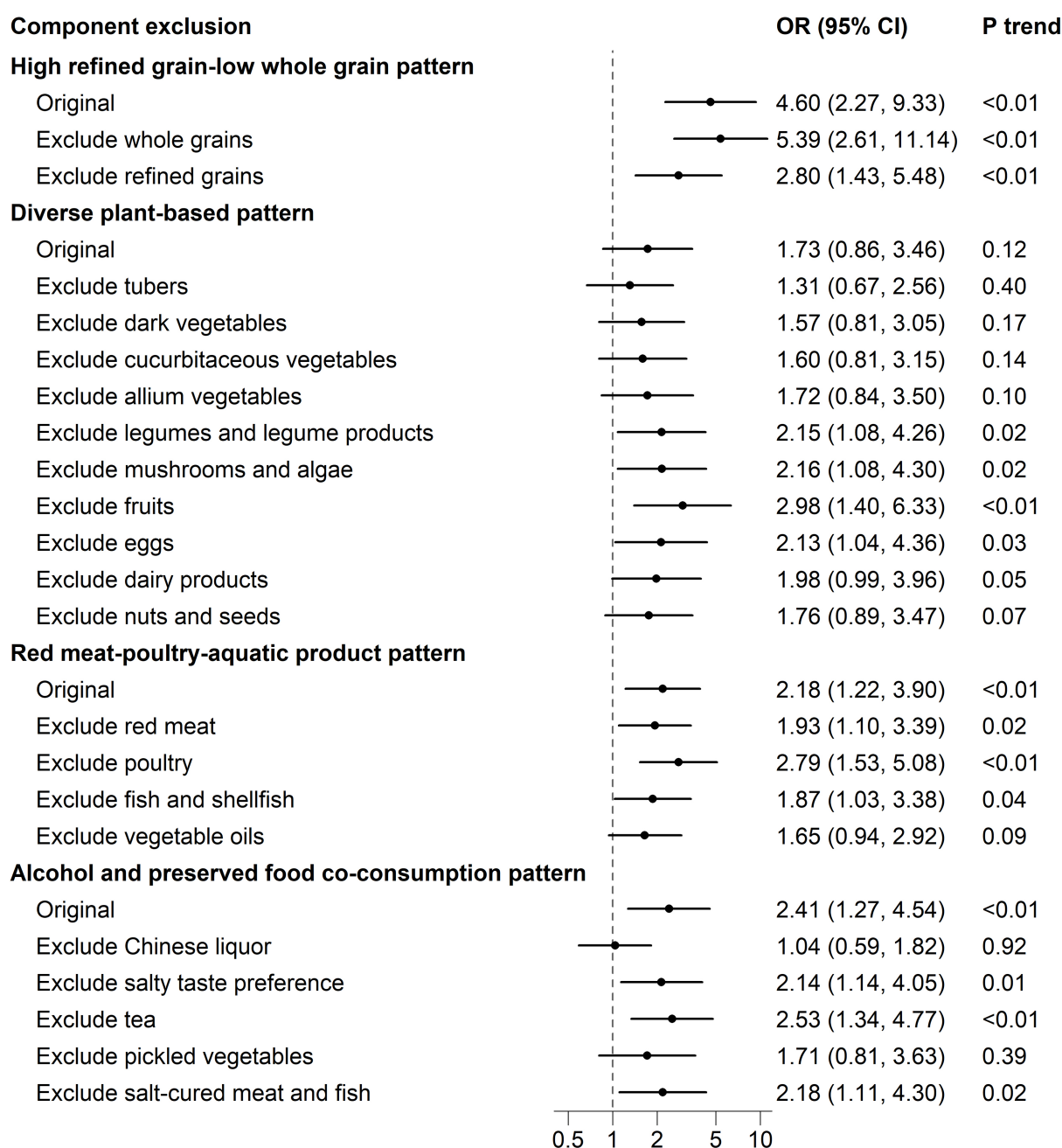


Figure 2. Forest plot of the component-exclusion sensitivity analyses for the associations between dietary pattern scores and the risk of PLC in the 1:2 matched case-control study ($n = 831$). The model was adjusted for education (primary school or below, junior high school, senior high school or above), annual household per capita income (< 10,000 Yuan, 10,000-20,000 Yuan, > 20,000 Yuan), energy intake (kcal/day, continuous), marital status, smoking status (never, former or current), alcohol drinking (never or light, moderate, heavy), physical activity (continuous), BMI (continuous), diabetes, HBV infection at baseline. The component exclusion approach involved iteratively recalculating the pattern scores by setting the loading of the specified food group to zero while retaining the original weights for all other loadings. The trend tests were conducted using the median value of each tertile as a continuous variable. BMI: Body mass index; OR: odds ratio; CI: confidence interval; PLC: primary liver cancer; HBV: hepatitis B virus.

Subgroup analysis

In the HBV-negative subgroup, the associations between dietary patterns and PLC risk were generally consistent with the primary results. No statistical interaction between dietary pattern and HBV infection status was observed for any of the 4 dietary patterns [Supplementary Table 4].

DISCUSSION

This study, employing a two-stage HCA-PCA approach, identified 4 major dietary patterns among adults in Anhui and examined their associations with PLC risk. We found that the high-refined-grain-low whole-grain pattern, the alcohol-and-preserved-food co-consumption pattern, and the red-meat-poultry-aquatic-product pattern were independently associated with an increased PLC risk. These associations appeared to be primarily attributable to specific components (Chinese liquor and pickled vegetables in the alcohol-and-preserved-food co-consumption pattern; red meat and vegetable oils for the red-meat-poultry-aquatic product pattern) and to a relative absence of potentially protective components (whole grains in the refined grain pattern). Given the imbalance in HBV infection between cases and controls at baseline, subgroup analyses were conducted according to HBV status. The associations observed in the HBV-negative subgroup were consistent with the main results; no statistical interaction was detected. The HBV-positive subgroup included very few controls ($n = 29$); therefore, the corresponding estimates were considered insufficiently informative and were not presented due to limited statistical power. Overall, these findings suggest that dietary patterns characterised by higher intake of refined grains, alcohol, pickled vegetables, vegetable oils, and meat may be relevant to PLC risk in this population.

The high-refined-grain-low whole-grain pattern, characterised by high-refined-grain and low whole-grain intake, was associated with increased PLC risk. Numerous prospective studies, including data from the Nurses' Health Study and the Health Professionals Follow-up Study, have reported an inverse association between higher whole grain intake and the risk of PLC^[29,30]. In contrast, evidence on refined grains is more limited. The Women's Health Initiative Observational Study and Nurses' Health Study found a positive association between total refined grain intake and PLC risk^[31,32]. It is noteworthy that the dietary context in China differs substantially from that of Western populations. While whole grains in Western diets often include oats, barley, and whole wheat, the predominant staples in China, particularly in southern and eastern regions such as Anhui, remain refined rice and wheat products, with whole grain consumption being low^[33,34]. In this context, the observed pattern may reflect poorer carbohydrate quality and lower fiber intake, which may partly help explain the positive association observed in our study^[35-37]. These findings suggest that, in this specific dietary setting, reducing reliance on refined staple foods while increasing whole-grain intake may represent a relevant dietary strategy to improve overall carbohydrate quality and potentially lower PLC risk.

The alcohol and preserved food co-consumption pattern, characterised by higher intake of Chinese liquor and salt-cured foods, was associated with an increased risk of PLC in our study. Existing evidence from large-scale cohorts in China has consistently reported a positive dose-response association between alcohol consumption and PLC, suggesting alcohol as an important contributor to this pattern^[38]. Mechanistically, although alcohol-related and metabolic fatty liver disease arise from different initial triggers, they share downstream pathways including oxidative stress, chronic inflammation, fibrosis, and gut-liver axis dysfunction, all of which may contribute to hepatocarcinogenesis^[39]. Our sensitivity analyses showed that the association was no longer observed after excluding Chinese liquor or pickled vegetables from the pattern score. In contrast, exclusion of tea strengthened the association, consistent with previous meta-analyses reporting an inverse association between tea consumption and PLC risk^[40,41]. Finally, exclusion of salt-cured meat and fish, or a salty taste preference, attenuated the risk estimate. These findings are consistent with the epidemiological literature, in which preserved foods and salt intake have been associated with increased cancer risk. For instance, a case-control study in Guangdong, China, identified a positive association between a preserved food-related dietary pattern and PLC, and other studies have similarly reported elevated risks of ovarian cancer and nasopharyngeal carcinoma associated with high intake of preserved foods and salt in China, the United Kingdom, and Japan^[42-44]. Overall, these findings suggest that higher consumption of Chinese liquor and preserved foods, along with lower tea intake, may be associated with an increased risk of PLC in this population.

The red meat-poultry-aquatic product pattern, characterised by high intakes of red meat, poultry, fish, and shellfish, was associated with an increased risk of PLC in our study. This finding is consistent with results from a hospital-based case-control study in the United States, a case-control study in China, the prospective UK Biobank study, and a meta-analysis, all of which reported positive associations between meat-dominant dietary patterns and PLC risk^[11,42,45,46]. In our component-exclusion sensitivity analyses, the positive association persisted after excluding red meat, fish, and shellfish, but was attenuated and became non-significant after excluding vegetable oils; by contrast, the association was strengthened when poultry was excluded. These findings are partly consistent with evidence from the National Institutes of Health - American Association of Retired Persons (NIH-AARP) Diet and Health Study, in which replacing red meat with poultry was associated with a lower risk of PLC, whereas replacing poultry with red meat was associated with a higher risk^[47]. They are also consistent with findings from the Nurses' Health Study, suggesting an inverse association between unsaturated fatty acid intake and PLC risk^[48]. Although prior studies have often reported inverse or null associations for fish intake^[12,13,49], the association observed after excluding fish and shellfish from our analysis may reflect the internal composition and correlation structure of this dietary pattern in the Anhui population, rather than a direct influence of fish and shellfish. This observation highlights the importance of interpreting pattern-based associations in light of local dietary context and the correlation structure of co-consumed foods. This suggests that, in this specific Anhui dietary context, the adverse effects of higher red meat intake and lower vegetable oil-related pattern expression may partly offset any potential benefit associated with co-consumed poultry products. Our findings suggest the dietary practice of appropriately reducing red meat frequency and portion size, while encouraging substitution with poultry and increasing the intake of vegetable oils.

The diverse plant-based pattern, characterised by high intake of vegetables, legumes and legume products, fruits, and other plant-derived foods, was not associated with PLC risk in our study ($BF_{01} = 5.68$). This finding differs from many previous studies, suggesting that higher intakes of plant foods are generally associated with a lower risk of PLC. For example, a case-control study in Guangdong and a prospective cohort study in Shanghai reported inverse associations between vegetable-based dietary patterns and PLC^[15,42,50]. A case-control study from the United States found that the inverse association between a prudent dietary pattern and PLC risk was more evident among women^[45]. However, some studies from Shanghai have also suggested that certain vegetables may be positively associated with PLC risk^[51]. In our component-exclusion sensitivity analyses, a positive association with PLC risk emerged after excluding fruits, legumes and legume products, mushrooms and algae, or eggs, whereas the exclusion of other plant-based components did not change the association. These findings are consistent with previous research on fruits, legumes and legume products, mushrooms and algae, or eggs^[15,42,52-54]. Several features of our study may help reconcile this discrepancy. One possibility is that the observed association may be subject to survivor bias and protopathic bias^[55,56]. Additionally, local cooking practices in Anhui often involve high-temperature stir-fries with oil and salt, rather than consumption in raw or minimally processed forms. Such preparation methods may reduce the levels of heat-sensitive nutrients and alter the nutritional profile of vegetables^[4,6,57]. Studies in Hong Kong have identified stir-fried vegetables as a major dietary source of acrylamide^[58]. Furthermore, high-temperature cooking may also promote the formation of advanced glycation end products and other heat-induced compounds^[59]. Notably, our sensitivity analysis provides some support for this interpretation. Excluding fruits revealed a positive association with PLC risk. This suggests that raw, antioxidant-rich fruits may partly offset the observed association within the pattern. Given this, we may need to focus not only on increasing the intake of plant-based foods but also on how foods are prepared and paired in the real world. For instance, suggesting fruit intake, reducing added oil and salt during vegetable preparation, and adopting less intensive cooking methods when feasible may help align plant-based eating with risk reduction.

The primary strength of this study is the use of an FFQ that has been validated for reliability and validity, supporting the accuracy of the dietary assessment^[19,20]. Additionally, the inclusion of PLC cases from a multicenter dataset enhances the representativeness of the study population. However, several limitations should be acknowledged. First, although we enrolled newly diagnosed PLC cases and excluded individuals with recent changes in dietary habits, total energy intake still differed between cases and controls. Therefore, the inherent case-control design precludes causal inference and may be susceptible to reverse causality. Second, protopathic bias is a particular concern, especially regarding the diverse plant-based patterns. Third, residual confounding cannot be ruled out. In particular, we lacked direct biomarkers of aflatoxin B1 exposure, detailed residence-based exposure information, and HCV data for the PLC case group. Therefore, both the grain-related associations and potential etiologic heterogeneity by viral hepatitis status should be interpreted with caution. Fourth, the component-exclusion analyses were intended to describe dietary structure and assess robustness, and should not be interpreted as reflecting the independent association of individual food groups. Finally, because dietary structures differ across regions, the generalizability of our findings requires further validation in other populations. Prospective cohort studies incorporating multi-omics approaches are needed to clarify the underlying biological mechanisms.

In conclusion, this study identified several dietary patterns associated with PLC risk in the Anhui population. Our findings suggest that dietary patterns characterised by refined grains, red meat, and the co-consumption of Chinese liquor and preserved foods were associated with an increased risk of PLC. These results highlight the need to consider region-specific dietary habits and cultural contexts when understanding PLC etiology. Further validation in prospective cohort studies is warranted.

DECLARATIONS

Acknowledgements

We would like to thank the participants and staff of the Anhui Lifestyle Validation Study for their valuable contributions.

Authors' contributions

Had full access to all study data and took responsibility for the integrity of the data and the accuracy of the data analysis: Chang Y

Study concept and design: Yang W

Acquisition, analysis, or interpretation of data: Chang Y, Chen J, Zhang Z, Yang W

Drafting of the manuscript: Chang Y

Critical revision of the manuscript for important intellectual content: Chang Y, Chen J, Zhang Z, Yang W

Statistical analysis: Chang Y

Obtained funding: Yang W

Administrative, technical, or material support: Yang W

Study supervision: Yang W

All authors read and approved the final manuscript.

Availability of data and materials

Data described in the manuscript, code book, and analytic code will be made available upon request, pending application.

AI and AI-assisted tools statement

Not applicable.

Financial support and sponsorship

This work was supported by the Major Projects in Natural Science of Higher Education Institutions in Anhui Province (2024AH040104), National Natural Science Foundation of China (grant number 82373673),

Research Funds of Center for Big Data and Population Health of IHM (grant number JKS2022018), grants from Anhui Medical University (grant number 2021xkjT007), and Postgraduate Innovation Research and Practice Program of Anhui Medical University (grant numbers YJS20230104, YJS20230150), Nutrition Science Foundation of Chinese Nutrition Society (grant number CNS-NCL2025-98).

Conflict of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

The ethics review board of Anhui Medical University approved the study protocol (Protocol Number: 20200897; 81250626), and all participants provided informed consent before enrollment.

Consent for publication

Not applicable

Copyright

© The Author(s) 2026.

Supplementary Materials

[Supplementary Materials](#)

REFERENCES

1. Sun KX, Li L, Wang SM, et al. Cancer incidence and mortality across diverse geographical regions in China, 2024. *Zhonghua Zhong Liu Za Zhi.* 2026;48:401-12. (in Chinese). [DOI PubMed](#)
2. Wang HD, Dai D, Zhu Q, et al. Cancer incidence and mortality in Anhui cancer registration areas in 2019 and trends from 2015 to 2019. *China Cancer.* 2025;34:98-107. [DOI](#)
3. Mak LY, Liu K, Chirapongsathorn S, et al. Liver diseases and hepatocellular carcinoma in the Asia-Pacific region: burden, trends, challenges and future directions. *Nat Rev Gastroenterol Hepatol.* 2024;21:834-51. [DOI PubMed](#)
4. Zhao W, Jia Z, Shi H. Nutrition and dietary inflammatory indices of the eight major cuisines of China. *BMC Nutr.* 2025;11:63. [DOI PubMed PMC](#)
5. Shen D. Research on integration and innovation of Anhui food eco-tourism industry under the concept of green development. *IOP Conf Ser : Earth Environ Sci.* 2021;769:022005. [DOI](#)
6. Mo H, Yin S, Liu Y. The development of traditional food in tourist destinations from the perspective of dramaturgy. *Sustainability.* 2022;14:16900. [DOI](#)
7. Zhu H, Li P, Wang L, Huang Q, Xu B. Flavor profile of “Dao Ban Xiang” (a traditional dry-cured meat product in Chinese Huizhou cuisine) at different processing stages in winter and summer. *Food Sci Nutr.* 2023;11:2733-50. [DOI PubMed PMC](#)
8. Yan S, Liu H, Zhang J, Tong Q. *Lactobacillus delbrueckii* is the key functional microorganism of natural fermented *tofu* sour water involved in the traditional coagulation of Chinese Huizhou *Mao-tofu*. *LWT.* 2020;131:109706. [DOI](#)
9. Kunutsor SK, Lehoczi A, Laukkanen JA. Coffee consumption, cancer, and healthy aging: epidemiological evidence and underlying mechanisms. *Geroscience.* 2025;47:1517-55. [DOI PubMed PMC](#)
10. Fridén M, Warensjö Lemming E, Lind L, Vessby J, Rosqvist F, Risérus U. Substitutions of saturated fat intakes with other macronutrients and foods and risk of NAFLD cirrhosis and all-cause hepatocellular carcinoma: a prospective cohort study. *Am J Clin Nutr.* 2024;120:187-95. [DOI PubMed](#)
11. Farvid MS, Sidahmed E, Spence ND, Mante Angua K, Rosner BA, Barnett JB. Consumption of red meat and processed meat and cancer incidence: a systematic review and meta-analysis of prospective studies. *Eur J Epidemiol.* 2021;36:937-51. [DOI PubMed](#)
12. Ma Y, Yang W, Li T, et al. Meat intake and risk of hepatocellular carcinoma in two large US prospective cohorts of women and men. *Int J Epidemiol.* 2019;48:1863-71. [DOI PubMed PMC](#)
13. Fedirko V, Trichopolou A, Bamia C, et al. Consumption of fish and meats and risk of hepatocellular carcinoma: the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann Oncol.* 2013;24:2166-73. [DOI PubMed PMC](#)
14. Shu W, Liu L, Jiang J, Yao Q. Dietary patterns and hepatocellular carcinoma risk: a systematic review and meta-analysis of cohort and case-control studies. *Nutr Metab.* 2024;21:47. [DOI PubMed PMC](#)
15. Shen QM, Tuo JY, Li ZY, et al. Sex-specific impact of dietary patterns on liver cancer incidence: updated results from two population-based cohort studies in China. *Eur J Nutr.* 2024;63:1113-24. [DOI PubMed](#)

16. Gorst-Rasmussen A, Dahm CC, Dethlefsen C, Scheike T, Overvad K. Exploring dietary patterns by using the treelet transform. *Am J Epidemiol.* 2011;173:1097-104. [DOI PubMed](#)
17. Wright DM, McKenna G, Nugent A, Winning L, Linden GJ, Woodside JV. Association between diet and periodontitis: a cross-sectional study of 10,000 NHANES participants. *Am J Clin Nutr.* 2020;112:1485-91. [DOI PubMed](#)
18. Chen J, Cao H, Xu Y, et al. Is light-to-moderate alcohol drinking associated with the onset of metabolic dysfunction-associated steatotic liver disease in a Chinese cohort? *Am J Clin Nutr.* 2026;123:101144. [DOI PubMed](#)
19. Zhang J, Liang S, Sun G, et al. Validation of a food frequency questionnaire for measuring dietary intake of isoflavones and lignans using the method of triads. *Am J Clin Nutr.* 2025;121:1157-64. [DOI PubMed](#)
20. Li M, Li X, Yin K, et al. Reproducibility and validity of low-carbohydrate diet and low-fat diet scores derived from 24-hour dietary recalls. *Am J Epidemiol.* 2025;194:2724-32. [DOI PubMed](#)
21. Willett W. *Nutritional epidemiology.* 3rd ed. Oxford: Oxford University Press; 2012. [DOI](#)
22. Feng C, Wang H, Lu N, et al. Log-transformation and its implications for data analysis. *Shanghai Arch Psychiatry.* 2014;26:105-9. [DOI PubMed PMC](#)
23. Sun C, Zhang WS, Jiang CQ, et al. Association of Cantonese dietary patterns with mortality risk in older Chinese: a 16-year follow-up of a Guangzhou Biobank cohort study. *Food Funct.* 2024;15:4538-51. [DOI PubMed](#)
24. Vijapur SM, Vaughan LE, Awan N, DiSanto D, McKernan GP, Wagner AK. Treelet transform analysis to identify clusters of systemic inflammatory variance in a population with moderate-to-severe traumatic brain injury. *Brain Behav Immun.* 2021;95:45-60. [DOI PubMed PMC](#)
25. Lorenzo-seva U, ten Berge JMF. Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology.* 2006;2:57-64. [DOI](#)
26. Hennig C. Cluster-wise assessment of cluster stability. *Comput Stat Data Anal.* 2007;52:258-71. [DOI](#)
27. Castelló A, Lope V, Vioque J, et al. Reproducibility of data-driven dietary patterns in two groups of adult Spanish women from different studies. *Br J Nutr.* 2016;116:734-42. [DOI PubMed](#)
28. Zhang S, Yan Y, Zeng XF, et al. Associations of the EAT-Lancet reference diet with metabolic dysfunction-associated steatotic liver disease and its severity: a multicohort study. *Hepatology.* 2025;81:1583-94. [DOI](#)
29. Yang W, Ma Y, Liu Y, et al. Association of intake of whole grains and dietary fiber with risk of hepatocellular carcinoma in US adults. *JAMA Oncol.* 2019;5:879-86. [DOI PubMed PMC](#)
30. Liu X, Yang W, Petrick JL, et al. Higher intake of whole grains and dietary fiber are associated with lower risk of liver cancer and chronic liver disease mortality. *Nat Commun.* 2021;12:6388. [DOI PubMed PMC](#)
31. Zhang X, Zhao L, Christopher CN, et al. Association of dietary insulinemic and inflammatory potential with risk of liver cancer and chronic liver disease mortality in postmenopausal women: a prospective cohort study. *Am J Clin Nutr.* 2023;118:530-7. [DOI PubMed PMC](#)
32. Liu Y, Yang W, VoPham T, et al. Plant-based and animal-based low-carbohydrate diets and risk of hepatocellular carcinoma among US men and women. *Hepatology.* 2021;73:175-85. [DOI PubMed](#)
33. Yu D, Zhao L, Zhao W. Status and trends in consumption of grains and dietary fiber among Chinese adults (1982-2015). *Nutr Rev.* 2020;78:43-53. [DOI PubMed](#)
34. Albertson A, Joshi N. Whole grain consumption and associations with body weight measures in the United States: results from NHANES 2009-10 and the new USDA Food Patterns Equivalents Database (810.23). *FASEB J.* 2014;28:810.23. [DOI](#)
35. Yu LX, Schwabe RF. The gut microbiome and liver cancer: mechanisms and clinical translation. *Nat Rev Gastroenterol Hepatol.* 2017;14:527-39. [DOI PubMed PMC](#)
36. Pollak M. Insulin and insulin-like growth factor signalling in neoplasia. *Nat Rev Cancer.* 2008;8:915-28. [DOI PubMed](#)
37. Rossi M, Lipworth L, Maso LD, et al. Dietary glycemic load and hepatocellular carcinoma with or without chronic hepatitis infection. *Ann Oncol.* 2009;20:1736-40. [DOI](#)
38. Im PK, Millwood IY, Kartsonaki C, et al.; China Kadoorie Biobank (CKB) Collaborative Group. Alcohol drinking and risks of total and site-specific cancers in China: a 10-year prospective study of 0.5 million adults. *Int J Cancer.* 2021;149:522-34. [DOI PubMed PMC](#)
39. Tarantino G, Citro V. What are the common downstream molecular events between alcoholic and nonalcoholic fatty liver? *Lipids Health Dis.* 2024;23:41. [DOI PubMed PMC](#)
40. Yu J, Liang D, Li J, et al. Coffee, green tea intake, and the risk of hepatocellular carcinoma: a systematic review and meta-analysis of observational studies. *Nutr Cancer.* 2023;75:1295-308. [DOI PubMed](#)
41. Kim TL, Jeong GH, Yang JW, et al. Tea consumption and risk of cancer: an umbrella review and meta-analysis of observational studies. *Adv Nutr.* 2020;11:1437-52. [DOI PubMed PMC](#)
42. Lan QY, Liao GC, Zhou RF, et al. Dietary patterns and primary liver cancer in Chinese adults: a case-control study. *Oncotarget.* 2018;9:27872-81. [DOI PubMed PMC](#)

43. Kliemann N, Rauber F, Bertazzi Levy R, et al. Food processing and cancer risk in Europe: results from the prospective EPIC cohort study. *Lancet Planet Health.* 2023;7:e219-32. DOI PubMed PMC
44. Zhang S, Huo Z, Borné Y, Sonestedt E, Qi L. Adding salt to foods and risk of metabolic dysfunction-associated steatotic liver disease and other chronic liver diseases. *Eur J Nutr.* 2025;64:224. DOI PubMed PMC
45. Moussa I, Day RS, Li R, et al. Dietary patterns and hepatocellular carcinoma risk among US adults. *Nutrients.* 2021;13:2011. DOI PubMed PMC
46. Guo W, Ge X, Lu J, et al. Diet and risk of non-alcoholic fatty liver disease, cirrhosis, and liver cancer: a large prospective cohort study in UK biobank. *Nutrients.* 2022;14:5335. DOI PubMed PMC
47. Wahl D, Loftfield E, Zouiouich S, et al. Meat consumption and risk of hepatobiliary cancers in the National Institutes of Health-AARP Diet and Health Study. *JNCI Cancer Spectr.* 2025;9:pkaf068. DOI PubMed PMC
48. Yang W, Sui J, Ma Y, et al. Dietary fat intake and risk of hepatocellular carcinoma in two large prospective cohort studies (FS13-07-19). *Curr Dev Nutr.* 2019;3:nzz030.FS13-07. DOI
49. Wang CR, Cai D, He K, et al. Red meat, poultry, and fish consumption and the risk of liver cancer: a prospective cohort study of 0.5 million chinese adults. *Cancer Epidemiol Biomarkers Prev.* 2025;34:412-9. DOI PubMed
50. Zhang W, Xiang YB, Li HL, et al. Vegetable-based dietary pattern and liver cancer risk: results from the Shanghai women's and men's health studies. *Cancer Sci.* 2013;104:1353-61. DOI PubMed PMC
51. Tuo JY, Shen QM, Li ZY, et al. A diet-wide association study for liver cancer risk: findings from a prospective cohort study in Chinese women. *J Nutr Sci.* 2024;13:e95. DOI PubMed PMC
52. Abe SK, Sawada N, Ishihara J, et al.; JPHC Study Group. Comparison between the impact of fermented and unfermented soy intake on the risk of liver cancer: the JPHC Study. *Eur J Nutr.* 2021;60:1389-401. DOI PubMed
53. Wang X, Liu Y. Anti-tumor activity of the water-soluble protein extract of tricholoma mongolicum (agaricomycetes) in H22 tumor-bearing mice and the possible underlying mechanism. *Int J Med Mushrooms.* 2023;25:75-85. DOI PubMed
54. Mahmoud YI, Shehata AMM, Fares NH, Mahmoud AA. Spirulina inhibits hepatocellular carcinoma through activating p53 and apoptosis and suppressing oxidative stress and angiogenesis. *Life Sci.* 2021;265:118827. DOI PubMed
55. Miller DP, Gomberg-Maitland M, Humbert M. Survivor bias and risk assessment. *Eur Respir J.* 2012;40:530-2. DOI PubMed
56. Horwitz RI, Feinstein AR. The problem of "protopathic bias" in case-control studies. *Am J Med.* 1980;68:255-8. DOI PubMed
57. Razzak A, Mahjabin T, Khan MRM, Hossain M, Sadia U, Zzaman W. Effect of cooking methods on the nutritional quality of selected vegetables at Sylhet City. *Heliyon.* 2023;9:e21709. DOI PubMed PMC
58. Wong WW, Chung SW, Lam CH, Ho YY, Xiao Y. Dietary exposure of Hong Kong adults to acrylamide: results of the first Hong Kong Total Diet Study. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* 2014;31:799-805. DOI PubMed
59. Uribarri J, Woodruff S, Goodman S, et al. Advanced glycation end products in foods and a practical guide to their reduction in the diet. *J Am Diet Assoc.* 2010;110:911-16.e12. DOI PubMed PMC

Disclaimer/Publisher's Note: All statements, opinions, and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect those of OAE and/or the editor(s). OAE and/or the editor(s) disclaim any responsibility for harm to persons or property resulting from the use of any ideas, methods, instructions, or products mentioned in the content.



© The Author(s) 2026. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.