

RESEARCH

Open Access



# The mediating roles of anthropo-metabolic biomarkers on the association between beverage consumption and breast cancer risk

Xiaoyi Lin<sup>1,6</sup>, Boheng Liang<sup>2</sup>, Tai Hing Lam<sup>3,4,6</sup>, Kar Keung Cheng<sup>5</sup>, Weisen Zhang<sup>4,6</sup> and Lin Xu<sup>1,3,5,6\*</sup>

## Abstract

**Background** Breast cancer (BC) is the most common malignancy in women, yet the role of beverage consumption in BC risk remains unclear. Additionally, the contribution of anthropo-metabolic biomarkers as mediators is unknown, limiting the development of effective prevention strategies.

**Methods** This study included 13,567 participants from the Guangzhou Biobank Cohort Study (GBCS), where beverage consumption was assessed at baseline using a food frequency questionnaire. BC cases were identified through cancer registry linkage over a mean follow-up of 14.8 years. Mendelian randomization (MR) analyses were performed to evaluate the causal effects of beverage consumption on BC risk, with a two-step MR approach used to estimate mediation effects.

**Results** During follow-up, 243 BC cases were identified. Weekly consumption of  $\geq 1$  portion of sugar sweetened beverages (SSB), versus  $< 1$  portion, was significantly associated with a higher risk of BC (hazard ratio [HR] 1.58, 95% confidence interval [CI] 1.12–2.23). This association was partly mediated by body mass index (proportion mediated [PM] 4.2%, 95% CI 0.9–17.1%) and uric acid (PM 18.8%, 95% CI 1.5–77.5%). Weekly consumption of  $> 6$  portions of dairy-based milk was associated with a non-significantly higher BC risk (HR 1.41, 95% CI 0.99–2.03), while 3–6 portions of soy milk were associated with a lower BC risk (HR 0.31, 95% CI 0.10–0.98). No significant associations were found for pure fruit juice, coffee, tea, or alcoholic drinks. MR analyses supported the detrimental effect of SSB on BC risk, with high-density lipoprotein cholesterol, polyunsaturated fatty acids to total fatty acids (TFAs) ratio, and omega-6 fatty acids to TFAs ratio mediating 2.44%, 2.73%, and 3.53% of the association, respectively.

**Conclusion** This study suggested that SSB consumption was a risk factor for BC and identified key anthropo-metabolic biomarkers mediating this relationship. Reducing SSB consumption and addressing associated metabolic pathways may offer effective strategies for BC prevention.

**Keywords** Beverage consumption, Breast cancer prevention, Anthro-metabolic biomarkers, Mediation, Sugar sweetened beverages

\*Correspondence:

Lin Xu

xulin27@mail.sysu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Introduction

The global beverage consumption has increased over the past three decades [1–3], driven by its contribution to nutrient and energy intake, and hedonic appeal. This dietary trend has become a priority concern due to its association with adverse health outcomes, including mortality, cardiovascular disease, and cancer [4–6]. In China, the per capita beverage consumption is approximately 120 kg annually, which contributes to a high disease burden, with southern China being one of the regions most affected [7]. Patterns of beverage consumption vary across population subgroups, with women generally consuming fewer sugar sweetened beverages (SSB) and alcoholic drinks, but more pure fruit juice (PFJ) and milk than men [1, 8].

Breast cancer (BC) is the most common malignancy among women globally, accounting for nearly one-third of all newly diagnosed female cancers [9]. China accounts for about 18% of global BC cases [10, 11], driven by a rising incidence and an aging population [12]. Identifying modifiable risk factors, such as dietary habits including beverage consumption, is essential for developing effective prevention strategies. Previous studies showed that higher SSB consumption was associated with a higher risk of BC [13, 14], and alcohol consumption is an established BC risk factor [15]. However, evidence for other beverages, including milk [16, 17], PFJ [5, 13], tea [15, 18] and coffee [19] remains inconsistent, with some studies suggesting potential risks or protective effects, while others find no association.

The association between beverage consumption and BC risk may be mediated through metabolic dysregulation, including obesity, insulin resistance, and altered lipid metabolism, all of which can promote carcinogenesis [20–22]. Additionally, the anti-inflammatory and antioxidant properties of certain beverages may modulate oxidative stress, systemic inflammation, and hormonal pathways, which are key contributors to BC pathogenesis [23, 24]. However, no studies have specifically examined biological mediators that might explain these associations or could be targeted for intervention.

To address these gaps, we conducted a prospective cohort study and Mendelian randomization (MR) analysis, accompanied by mediation analyses to investigate the associations between beverage consumption and BC risk and to identify anthropo-metabolic biomarkers mediating these relationships. MR analysis uses genetic variants as instrumental variables, providing estimates that are less susceptible to confounding or reverse causation compared to conventional observation studies [25]. Mediation analysis additionally explores the intermediate variables through which an exposure influences an outcome [26].

## Methods

### Observational study

#### Study population

The Guangzhou Biobank Cohort Study (GBCS) is a three-way collaboration of the Guangzhou Twelfth People's Hospital and the Universities of Hong Kong and Birmingham. Details have previously been described [27]. In brief, this longitudinal cohort consists of participants from the “Guangzhou Health and Happiness Association for the Respectable Elders” (GHHARE), a community social and welfare organization. The GHHARE included about 7% of Guangzhou permanent residents aged 50 years or older in all 10 districts of Guangzhou, the capital city of Guangdong province in southern China. GHHARE participants were eligible if they were ambulatory, capable of providing informed consent, and not undergoing treatments for life-threatening conditions. Only those recruited in phases 1 and 2 (2003–2006) were included in the current study, as the Food Frequency Questionnaire (FFQ) was shortened in phase 3 (2006–2008), limiting dietary exposure assessment.

Baseline information was collected using a computer-based standardized questionnaire by face-to-face interviews. The reproducibility of the questionnaire responses was tested by re-interviewing 200 randomly selected participants after a 1-month interval, which generated satisfactory results [27]. All laboratory analyses were performed on fresh blood samples in the Clinical Laboratory of the Guangzhou Twelfth People's Hospital using standardized, automated, well-documented methodologies [27]. Physical examination was done by trained nurses in the hospital using standard protocols [27]. The Guangzhou Medical Ethics Committee of the Chinese Medical Association approved the study, and all participants provided written, informed consent before participation.

#### Exposures, outcome, and potential mediators

Baseline beverage consumption was assessed using a 300-item validated FFQ [28]. The FFQ included 19 commonly consumed beverages in Southern China, with intake calculated as the product of the number of portions per occasion and weekly frequency, expressed in portions per week (one portion = 250 mL). Beverages were categorized into seven types: dairy-based milk, soy milk, SSB, PFJ, coffee, tea, and alcoholic drinks (Table S1). Dairy-based milk and soy milk consumption were classified into four groups (<1, 1–2, 3–6, >6 portions/week), and SSB, PFJ, coffee, tea, and alcoholic drinks were classified into two groups (<1, ≥1 portion/week) due to the relatively small number of high consumers.

Participants were followed until December 31, 2020, through linkage with the cancer registry and death registry of the Guangzhou Center for Disease Control and

Prevention using their unique resident identity card numbers. BC cases were identified and coded as “C50” based on the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) by trained clinical coding officers at each hospital.

To investigate potential mechanisms, 12 anthropo-metabolic parameters previously associated with BC risk were assessed as mediators [29–33]. Anthropometric parameters included body mass index (BMI), waist circumference, and waist-to-hip ratio. BMI was calculated by dividing the weight (kg) by height squared ( $m^2$ ). Serum metabolic parameters included fasting glucose, lipids (total cholesterol, low-density lipoprotein cholesterol [LDL-C], high-density lipoprotein cholesterol [HDL-C], triglycerides), bilirubin, blood urea nitrogen, creatinine, and uric acid.

## Mendelian randomization

### Instrument selection

Summary-level genome-wide association studies (GWAS) data for beverage intake (dairy-based milk, soy milk, SSB, PFJ, coffee, tea, and alcoholic drinks) were obtained from the UK Biobank, which used web-based 24-h recall questionnaire and touchscreen questionnaire at assessment center visit [34, 35]. Genetic instruments for BC were obtained from a meta-analysis of 67 GWAS studies including 122,977 cases and 105,974 controls [36]. Genetic instruments of 25 anthropo-metabolic markers, previously identified as BC risk predictors [29–33], were selected for mediation analyses [37–43]. Details about the GWAS data sources were summarized in Table S2.

Genetic instruments for each exposure and mediator were selected based on a genome-wide significant threshold ( $P < 5e-8$ ), and independence criteria (linkage disequilibrium  $r^2 < 0.001$  within 10,000 kb). The thresholds for dairy-based milk, soy milk, SSB, PFJ, coffee, and tea were set at  $P < 5e-6$  [34], due to the limited number of single nucleotide polymorphisms (SNPs) meeting the stringent criterion.

## Statistical analysis

### Observational study

Baseline characteristics by BC status were compared using the t-test for continuous variables and Pearson's Chi-squared test for categorical variables. Potential confounders included age, socioeconomic position (education level, occupation, annual personal income), behavioral factors (smoking status, alcohol use, physical activity), reproductive factors (age at menarche and menopause, parity and breastfeeding history), personal and family medical history (oral contraceptive use, hormone replacement therapy, self-reported health status, and family history of BC), and daily dietary energy intake.

Beverage consumption was analyzed both as a categorical and a continuous variable. Cox proportional hazards regression was used to estimate the association between beverage consumption and BC risk, yielding crude and adjusted hazard ratios (HRs) and 95% confidence intervals (CIs). The proportional hazards assumption was assessed using the Schoenfeld residual test, with no violation detected. Stratified analyses were conducted by age ( $< 60 / \geq 60$  years) and menopausal status. Sensitivity analyses were performed. First, to minimize reverse causality, analyses were repeated after excluding BC cases or deaths occurring within the first year of follow-up. Second, to account for competing risks, proportional subdistribution hazards regression was used to estimate the subdistribution HRs [44]. Third, four partially adjusted models were constructed, each accounting for a specific set of potential confounders including socioeconomic position, behavioral factors, reproductive factors, and personal and family medical history, to assess the influence of these various confounder categories.

Linear regression was used to estimate the associations between beverage consumption and anthropo-metabolic parameters, with the assumptions of linearity, normality of residuals, homoscedasticity, and independence being tested and verified. Significant markers were subsequently included in mediation analyses. Mediation proportions were estimated using the difference method by comparing regression models with and without the inclusion of the mediator [45, 46].

### Mendelian randomization

Instrument strength was evaluated using the F-statistic, with a value  $> 10$  considered sufficient to minimize weak instrument bias. Harmonization of genetic effects was performed prior to analysis to ensure consistency in directionality between exposure and outcome associations. For each independent SNP ( $r^2 < 0.001$ ), the Wald ratio was calculated as the SNP-outcome association divided by the SNP-exposure association. These estimations were combined using inverse-variance weighted (IVW) with random effects independent [47]. Heterogeneity among SNPs was assessed using Cochran's Q test [48]. The weighted median, MR-Egger, and MR pleiotropy residual sum and outlier (MR-PRESSO) methods were performed to validate the robustness of the IVW results based on different assumptions.

Univariable MR estimated the total effect of the exposure on the outcome. Then, two-step MR was used to estimate the indirect effect of the exposure on the outcome through a mediator [26]. Univariable MR was first used to estimate the causal effect of the exposure on the mediator and the causal effect of the mediator on the outcome. The indirect effect was calculated by multiplying

these two estimates, following the product-of-coefficients method. The delta method was used to approximate the CIs of mediation effects [49]. Negative mediation proportions were truncated at 0%, as this is the minimum threshold for determining mediation. IVW was used as the primary method for estimating mediating effects.

Statistical significance was defined as a two-sided  $P$ -value  $< 0.05$ . Observational mediation analyses were conducted using the %MEDIATE Macro in SAS (version 9.4; SAS Institute, Cary, NC). All other statistical analyses were conducted using R software (version 4.3.1).

## Results

### Baseline characteristics

Among GBCS participants (phase 1–2) who completed the FFQ, participants with a self-reported or documented cancer diagnosis, those lost to follow-up, male participants, and those with missing data on confounders were excluded, yielding an analytical sample of 13,567 participants (Figure S1). During an average follow-up period of 14.8 years with 201,277 person-years, 243 BC cases were identified. Table 1 shows that participants who developed BC were younger, had higher education levels and annual personal income, and a greater proportion of non-manual workers (all  $P$ -values  $< 0.05$ ). Additionally, the BC group also had a lower prevalence of active physical activity and a higher proportion of consuming  $\geq 1$  portion of SSB per week (all  $P$ -values  $< 0.05$ ).

### Association between beverage consumption and BC risk

Table 2 shows that, after adjusting for 16 potential confounders, the consumption of  $> 6$  portions of dairy-based milk per week was marginally associated with an increased risk of BC (HR 1.41, 95% CI 0.99–2.03,  $P = 0.06$ ), compared with  $< 1$  portion per week. In contrast, participants consuming 3–6 portions of soy milk per week showed a marginally lower risk of BC (HR 0.31, 95% CI 0.10–0.98,  $P = 0.047$ ). For SSB, consumption of  $\geq 1$  portion per week was significantly associated with an increased risk of BC (HR 1.58, 95% CI 1.12–2.23,  $P = 0.009$ ), with each additional portion associated with a 5% higher risk of BC (95% CI 1.03–1.07,  $P < 0.001$ ). No significant associations were observed between BC risk and the consumption of PFJ, coffee, tea, or alcoholic drinks.

A significant interaction between SSB consumption and age was observed, with the positive association between SSB consumption and BC risk being statistically significant only in participants younger than 60 years (Table S3). Stratification by menopausal status showed that this association was significant only in postmenopausal women (Table S4). In sensitivity analyses, the positive association between SSB consumption and BC risk

remained after excluding BC cases or deaths occurring within the first year of follow-up (Table S5). Results from the competing risk analysis were consistent with the main analysis (Table S6). The association of SSB consumption with BC risk remained across all partially adjusted models (Table S7).

### Mediation analysis

For the mediation analysis, we further restricted to 13,359 participants with complete data on BMI, waist circumference, waist-to-hip ratio, fasting glucose, serum total cholesterol, LDL-C, and HDL-C, serum total triglycerides, and to 6,855 participants with additional data on total bilirubin, blood urea nitrogen, serum creatinine, and uric acid (Figure S1). Table S8 shows that compared with  $< 1$  portion/week, consumption of 3–6 portions of soy milk per week was significantly associated with higher levels of BMI and waist circumference (all  $P$ -values  $< 0.01$ ). Consumption of  $\geq 1$  portion of SSB was significantly associated with higher levels of BMI, LDL-C, and uric acid, but lower levels of fasting glucose (all  $P$ -values  $< 0.05$ ). In Fig. 1, mediation analysis identified BMI and uric acid as mediators of the association of SSB consumption with higher BC risk, with a mediation proportion of 4.2% (95% CI 0.9–17.1%,  $P = 0.007$ ) and 18.8% (95% CI 1.5–77.5%,  $P = 0.02$ ), respectively, and a total mediation effect of 22.4% (95% CI 1.7–83.0%,  $P = 0.006$ ). No significant mediation was observed for the association between soy milk consumption and BC risk (Table S9).

### Mendelian randomization

#### Causal effect of beverage consumption on BC

A total of 20, 27, 14, 9, 81, 86, and 77 SNPs were selected as instruments for dairy-based milk, soy milk, SSB, PFJ, coffee, tea, and alcoholic drinks consumption, respectively. The average F-statistics for these instruments were 24.3, 23.2, 24.1, 21.8, 44.7, 31.8 and 62.3. Table S10 summarizes the genetic instruments used for beverage consumption.

Table 3 shows that higher genetically determined SSB consumption was associated with an increased BC risk (odds ratio [OR] 3.52, 95% CI 1.06–11.70,  $P = 0.04$ ). The MR-PRESSO method identified two SNPs as influential outliers, and the outlier-corrected estimates remained statistically significant (OR 1.98, 95% CI 1.48–2.66,  $P < 0.001$ ). The direction of the association was consistent across all four MR methods, although weighted median and MR-Egger results had wider CIs. No evidence for directional horizontal pleiotropy was detected ( $P$  for MR-Egger intercept = 0.75). However, genetically determined consumption of dairy-based milk, soy milk, PFJ, coffee, tea, and alcoholic drinks were not associated

**Table 1** Baseline characteristics of participants in the Guangzhou Biobank Cohort Study

Characteristics	Breast cancer		P-value
	No (N = 13,324)	Yes (N = 243)	
Age, years (mean ± SD)	61.92 ± 6.63	61.05 ± 6.63	0.04
<b>Socioeconomic position</b>			
Education level			< 0.001
Primary or below	6,900 (52%)	92 (38%)	
Middle school or above	6,424 (48%)	151 (62%)	
Occupation			< 0.001
Manual	8,817 (66%)	133 (55%)	
Non-manual	2,574 (19%)	56 (23%)	
Others	1,933 (15%)	54 (22%)	
Personal income, RMB/year			0.01
< 10,000	5,726 (43%)	87 (36%)	
≥ 10,000	6,935 (52%)	149 (61%)	
Not reported	663 (5.0%)	7 (2.9%)	
<b>Behavioral factors</b>			
Smoking status			0.40
Never	12,798 (96%)	236 (97%)	
Former or current	526 (3.9%)	7 (2.9%)	
Alcohol use			0.99
Never	11,949 (90%)	218 (90%)	
Former or current	1,375 (10%)	25 (10%)	
Physical activity			0.01
Inactive/Moderate	7,274 (55%)	152 (63%)	
Active	6,050 (45%)	91 (37%)	
<b>Reproductive factors</b>			
Age at menarche, years			0.08
> 12	12,303 (92%)	217 (89%)	
≤ 12	1,021 (7.7%)	26 (11%)	
Age at menopause, years			0.30
< 45	1,187 (8.9%)	17 (7.0%)	
≥ 45	12,137 (91%)	226 (93%)	
Parity, no	1,312 (9.8%)	15 (6.2%)	0.06
Breastfeeding history, no	1,733 (13%)	26 (11%)	0.29
<b>Personal and family medical history</b>			
Oral contraceptive use, yes	2,299 (17%)	46 (19%)	0.49
Hormone replacement therapy, yes	145 (1.1%)	4 (1.6%)	0.61
Self-reported health status, poor	2,566 (19%)	59 (24%)	0.05
Family history of breast cancer, yes	105 (0.8%)	4 (1.6%)	0.26
<b>Dietary factors</b>			
Daily dietary energy intake, kcal (mean ± SD)	1,757.74 ± 497.76	1,769.28 ± 526.81	0.74
Dairy-based milk			0.05
< 1 portion/week	9,222 (69%)	151 (62%)	
1–2 portions/week	1,163 (8.7%)	25 (10%)	
3–6 portions/week	1,469 (11%)	28 (12%)	
> 6 portions/week	1,470 (11%)	39 (16%)	
Soy milk			0.18
< 1 portion/week	11,308 (85%)	210 (86%)	
1–2 portions/week	1,355 (10%)	28 (12%)	
3–6 portions/week	493 (3.7%)	3 (1.2%)	



**Table 1** (continued)

Characteristics	Breast cancer		P-value
	No (N = 13,324)	Yes (N = 243)	
> 6 portions/week	168 (1.3%)	2 (0.8%)	0.003
Sugar sweetened beverages			
< 1 portion/week	11,869 (89%)	202 (83%)	
≥ 1 portion/week	1,455 (11%)	41 (17%)	0.44
Pure fruit juice			
< 1 portion/week	12,989 (97%)	235 (97%)	
≥ 1 portion/week	335 (2.5%)	8 (3.3%)	0.39
Coffee			
< 1 portion/week	13,046 (98%)	236 (97%)	
≥ 1 portion/week	278 (2.1%)	7 (2.9%)	0.92
Tea			
< 1 portion/week	9,611 (72%)	176 (72%)	
≥ 1 portion/week	3,713 (28%)	67 (28%)	0.40
Alcoholic drinks			
< 1 portion/week	13,045 (98%)	236 (97%)	
≥ 1 portion/week	279 (2.1%)	7 (2.9%)	

Abbreviations: SD, standard deviation. 1 portion = 250 ml

with BC risk in any MR analysis. Regarding BC subtypes (Table S11), genetically determined SSB consumption was consistently associated with an increased risk of estrogen receptor (ER)-negative BC (OR 5.69, 95% CI 1.22–26.63,  $P=0.03$ ), and was marginally associated with an increased risk of ER-positive BC (OR 3.07, 95% CI 0.94–9.98,  $P=0.06$ ).

#### Mediation analysis

Among 25 anthropo-metabolic biomarkers (Table S12), genetically determined higher consumption of SSB was causally associated with HDL-C, the ratio of polyunsaturated fatty acids (PUFAs) to total fatty acids (TFAs), the ratio of omega-6 fatty acids to TFAs, and lower total triglycerides (all  $P$ -values < 0.05). No evidence of directional horizontal pleiotropy was observed (all  $P$ -values for MR-Egger intercept  $\geq 0.33$ ). For these four potential mediators (Table S13), genetically determined HDL-C, ratio of PUFAs to TFAs, and ratio of omega-6 fatty acids to TFAs were positively associated with an increased risk of BC (all  $P$ -values < 0.05). Instrumental validity test confirmed sufficient instrument strength for these biomarkers (all  $F$ -statistics > 110), with no evidence for horizontal pleiotropy (all  $P$  for MR-Egger intercept  $\geq 0.06$ ). Following mediator selection, two-step MR analysis showed that HDL-C, the ratio of PUFAs to TFAs, and the ratio of omega-6 fatty acids to TFAs mediated 2.44%, 2.73%, and 3.53% of the total effect of SSB consumption on BC risk, respectively (Fig. 2).

#### Discussion

This study, triangulating evidence from observational and MR analyses, showed that higher SSB consumption was causally associated with an increased BC risk. Mediation analyses identified key anthropo-metabolic biomarkers, including BMI, uric acid, HDL-C, ratio of PUFAs to TFAs, and ratio of omega-6 fatty acids to TFAs as mediators of this association. Notably, these mediators explained a substantial proportion of the total effect of SSBs on BC risk, highlighting potential biological pathways that linked dietary habits to cancer risk. Furthermore, while dairy-based milk and soy milk showed marginal associations with BC risk, no significant associations were observed for PFJ, coffee, tea, or alcohol consumption. These findings indicated the potential causal links between SSB consumption and BC risk, and suggested the importance of addressing specific metabolic mediators, such as BMI and uric acid, in developing targeted cancer prevention strategies.

#### Comparison with other studies

Our findings were consistent with the NutriNet-Santé prospective cohort, which identified a positive association between sugary drink consumption and BC risk in adults aged  $\geq 18$  years [5]. However, this study used an extensive dietary assessment approach, incorporating nearly 100 sugary drink items and repeated dietary measurements over time, which enhances dietary exposure assessment precision but may limit generalizability to other populations. Similarly, another prospective

**Table 2** Associations between beverage consumption and risk of breast cancer in the Guangzhou Biobank Cohort Study

	N	Incidence rate/1000 person-years	Crude HR (95% CI)	Adjusted HR (95% CI) <sup>a</sup>	P <sup>b</sup>
<b>Dairy-based milk</b>					
< 1 portion/week	9,373	1.08	1.00	1.00	
1–2 portions/week	1,188	1.44	1.33 (0.87, 2.03)	1.15 (0.75, 1.77)	0.51
3–6 portions/week	1,497	1.26	1.16 (0.77, 1.74)	0.98 (0.65, 1.48)	0.94
> 6 portions/week	1,509	1.76	1.62 (1.14, 2.31)	1.41 (0.99, 2.03)	0.06
per 1 portion			1.05 (1.01, 1.09)	1.03 (0.99, 1.08)	0.15
<b>Soy milk</b>					
< 1 portion/week	11,518	1.23	1.00	1.00	
1–2 portions/week	1,383	1.35	1.09 (0.74, 1.62)	1.04 (0.70, 1.54)	0.87
3–6 portions/week	496	0.39	0.32 (0.10, 0.99)	0.31 (0.10, 0.98)	0.047
> 6 portions/week	170	0.79	0.64 (0.16, 2.59)	0.63 (0.16, 2.54)	0.52
per 1 portion			0.89 (0.77, 1.03)	0.88 (0.76, 1.02)	0.09
<b>Sugar sweetened beverages</b>					
< 1 portion/week	12,071	1.13	1.00	1.00	
≥ 1 portion/week	1,496	1.88	1.67 (1.19, 2.33)	1.58 (1.12, 2.23)	0.009
per 1 portion			1.05 (1.03, 1.07)	1.05 (1.03, 1.07)	< 0.001
<b>Pure fruit juice</b>					
< 1 portion/week	13,224	1.20	1.00	1.00	
≥ 1 portion/week	343	1.57	1.31 (0.65, 2.65)	1.22 (0.60, 2.49)	0.57
per 1 portion			1.13 (0.95, 1.34)	1.13 (0.94, 1.36)	0.18
<b>Coffee</b>					
< 1 portion/week	13,282	1.20	1.00	1.00	
≥ 1 portion/week	285	1.69	1.42 (0.67, 3.01)	1.19 (0.56, 2.54)	0.65
per 1 portion			0.87 (0.56, 1.35)	0.82 (0.52, 1.29)	0.39
<b>Tea</b>					
< 1 portion/week	9,787	1.21	1.00	1.00	
≥ 1 portion/week	3,780	1.19	0.98 (0.74, 1.30)	1.02 (0.77, 1.35)	0.90
per 1 portion			1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	0.65
<b>Alcoholic drinks</b>					
< 1 portion/week	13,281	1.20	1.00	1.00	
≥ 1 portion/week	286	1.69	1.42 (0.67, 3.01)	1.22 (0.56, 2.62)	0.62
per 1 portion			1.00 (0.82, 1.22)	0.99 (0.73, 1.34)	0.93

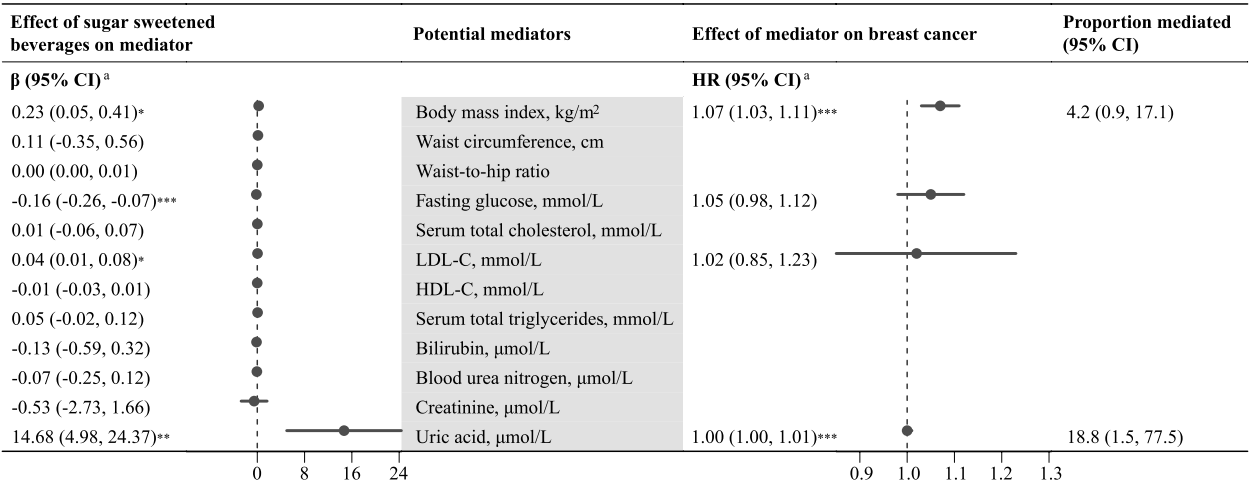
Abbreviations: HR hazard ratio, CI confidence interval. 1 portion = 250 ml

<sup>a</sup> Adjusted for age, education level, occupation, annual personal income, smoking status, alcohol use, physical activity, age at menarche and menopause, parity and breastfeeding history, oral contraceptive use, hormone replacement therapy, self-reported health status, family history of breast cancer, and daily dietary energy intake

<sup>b</sup> P-value was for the adjusted model

study conducted on middle-aged university graduates reported a positive association between sugary drink consumption and BC risk, but only in postmenopausal women, not in premenopausal women [50]. In contrast, the Nurses' Health Study and the Canadian Study of Diet, Lifestyle, and Health reported no association between SSB consumption and BC risk when using < 1/month or none as the reference group [51, 52]. These discrepancies may reflect differences in study design, including the definitions of SSB exposure and variation in population

characteristics. For example, the null findings in the Nurses' Health Study and the Canadian cohort may be attributed to the younger participant profiles with mean ages < 50 years, which could lower the overall BC risk and attenuate the association with SSB consumption [51, 52]. Regarding milk consumption, no association with BC risk was reported in a cohort of over 60,000 female participants [53]. However, dairy milk intake at the 90th percentile was associated with an increased BC risk in a cohort of more than 50,000 women [54]. Furthermore,



**Fig. 1** Mediating effects of the association between sugar sweetened beverages and risk of breast cancer by anthropo-metabolic markers in the Guangzhou Biobank Cohort Study. Abbreviations: CI, confidence interval; HR, hazard ratio; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol. <sup>a</sup>Adjusted for age, education level, occupation, annual personal income, smoking status, alcohol use, physical activity, age at menarche and menopause, parity and breastfeeding history, oral contraceptive use, hormone replacement therapy, self-reported health status, family history of breast cancer, and daily dietary energy intake. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$

a previous MR study using a single SNP near the lactase gene locus as a proxy for milk consumption reported a positive association between milk consumption and BC risk [17]. These discrepancies may be due to variations in milk consumption classifications and genetic predispositions related to dairy digestion. Our findings on soy milk consumption were consistent with a study that analyzed soy milk intake as a continuous variable, which also found no significant linear association with BC risk [54]. However, that study reported that substituting dairy milk intake with soy milk was associated with a lower BC risk, suggesting a potential protective effect when replacing specific dietary components rather than from soy milk alone [54]. Our results for PFJ and tea consumption were also consistent with findings from large-scale cohort studies, which reported no significant associations with BC risk [5, 19]. For coffee, our findings were consistent with an American study that found no association between coffee intake and postmenopausal BC risk [55]. However, a European cohort study that differentiated between caffeinated and decaffeinated coffee suggested that higher caffeinated coffee intake may be associated with a lower risk of postmenopausal BC [19]. These discrepancies highlight the potential influence of coffee type and preparation methods on cancer risk, warranting further investigation. Regarding alcoholic drinks, our findings aligned with those of a Japanese study reporting no association between alcohol consumption and BC risk among Asian postmenopausal women [56]. Given the relatively low alcohol intake among Chinese women and the smaller sample size, our study may not have sufficient

statistical power to detect a potential association. Furthermore, the low prevalence of obesity in our population could attenuate the effects of alcohol-induced estrogen production from adipose tissue [56, 57].

**Mechanisms**

The association between SSB consumption and BC risk is most commonly attributed to obesity [20, 58], supported by our study which identified BMI as a potential mediator. Chronic SSB intake contributes to weight gain through excessive caloric intake and metabolic dysregulation, establishing a link between BMI and BC risk. Furthermore, uric acid emerged as an additional mediator. Fructose metabolism in SSB promotes ATP degradation and de novo purine synthesis, leading to elevated uric acid levels [59–61]. While evidence regarding the direct association between uric acid and BC risk remains inconclusive [62, 63], uric acid is a known inducer of oxidative stress, potentially driving the malignant transformation of breast cells [64, 65]. Additionally, uric acid may act as a secondary mediator, linking BMI to BC risk [33], as higher BMI can impair renal tubular excretion of uric acid in the context of SSB consumption [66]. In addition to BMI and uric acid, our MR analysis indicated a positive causal association of HDL-C with BC risk, consistent with previous study [67], However, this finding contrasted with observational evidence showing SSB consumption was associated with lower HDL-C levels [68]. This discrepancy underscores the complexity of lipid metabolism in relation to SSB intake and BC risk, highlighting the need for further investigation to clarify



**Table 3** Two-sample Mendelian randomization estimates for the causal associations between beverage consumption and overall breast cancer

Exposure	SNP	F statistic	Methods	OR (95% CI)	P	Cochran's Q statistic (I <sup>2</sup> )	MR-Egger Intercept (P)
Dairy-based milk	20	24.3	IVW	1.26 (0.75, 2.12)	0.38	14.54 (0.0%)	-0.002 (0.77)
			WM	1.06 (0.51, 2.20)	0.87		
			MR-Egger	1.63 (0.28, 9.47)	0.60		
			MR-PRESSO	-	-		
Soy milk	27	23.2	IVW	0.98 (0.92, 1.05)	0.59	35.76 (27.3%)	-0.002 (0.81)
			WM	0.96 (0.88, 1.05)	0.39		
			MR-Egger	1.01 (0.82, 1.24)	0.96		
			MR-PRESSO	-	-		
Sugar sweetened beverages	14	24.1	IVW	3.52 (1.06, 11.70)	0.04	85.10 (84.7%)	0.006 (0.75)
			WM	1.60 (0.81, 3.17)	0.18		
			MR-Egger	1.99 (0.05, 80.07)	0.72		
			MR-PRESSO <sup>a</sup>	1.98 (1.48, 2.66)	<0.001		
Pure fruit juice	9	21.8	IVW	3.07 (0.38, 24.94)	0.29	83.49 (90.4%)	0.030 (0.31)
			WM	0.99 (0.41, 2.40)	0.98		
			MR-Egger	0.24 (0.00, 36.10)	0.59		
			MR-PRESSO <sup>a</sup>	0.85 (0.51, 1.42)	0.55		
Coffee	81	44.7	IVW	0.87 (0.67, 1.13)	0.29	301.33 (73.5%)	-0.004 (0.28)
			WM	0.92 (0.71, 1.20)	0.54		
			MR-Egger	1.12 (0.66, 1.88)	0.68		
			MR-PRESSO <sup>a</sup>	0.84 (0.68, 1.04)	0.11		
Tea	86	31.8	IVW	1.07 (0.88, 1.31)	0.50	172.71 (50.8%)	-0.003 (0.35)
			WM	1.02 (0.81, 1.28)	0.87		
			MR-Egger	1.43 (0.76, 2.70)	0.27		
			MR-PRESSO <sup>a</sup>	0.92 (0.80, 1.06)	0.23		
Alcoholic drinks	77	62.3	IVW	1.02 (0.88, 1.19)	0.79	262.73 (71.1%)	0.001 (0.69)
			WM	0.95 (0.83, 1.09)	0.50		
			MR-Egger	0.98 (0.76, 1.26)	0.87		
			MR-PRESSO <sup>a</sup>	0.99 (0.89, 1.11)	0.91		

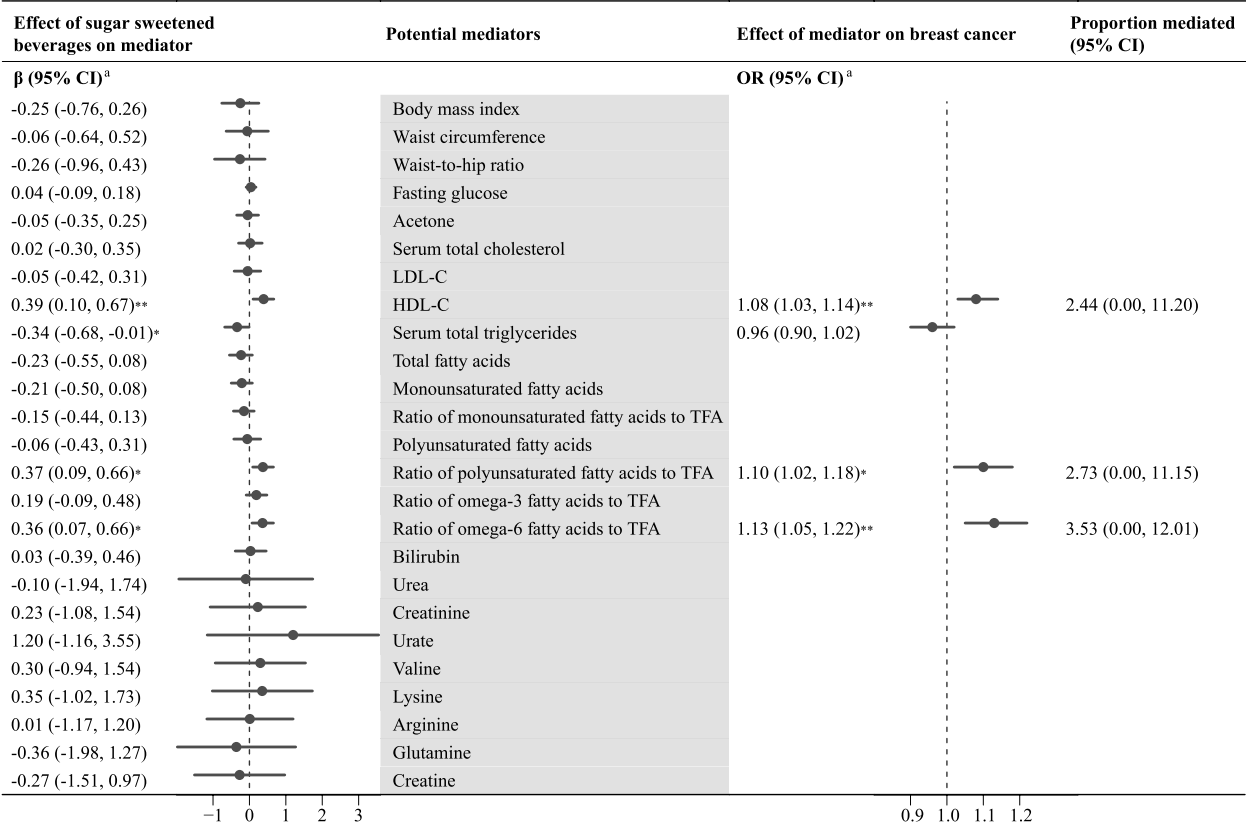
**Abbreviations:** SNP, single nucleotide polymorphism; IVW, inverse-variance weighted; WM, weighted median method; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier; OR, odds ratio; CI, confidence interval

<sup>a</sup> 2 SNPs were identified as influential outliers when consumption of sugar sweetened beverages was exposure: rs2472297, rs55872725; 1 SNP was identified as influential outliers when consumption of pure fruit juice was exposure: rs9972653; 7 SNPs were identified as influential outliers when consumption of coffee was exposure: rs10865548, rs111994577, rs2231142, rs2472297, rs2521501, rs57918684, rs9937053; 4 SNPs were identified as influential outliers when consumption of tea was exposure: rs11715828, rs1481012, rs2074551, rs2271961; 4 SNPs were identified as influential outliers when consumption of alcoholic drinks was exposure: rs12030672, rs2959005, rs56094641, rs62244890

these relationships. Alterations in fatty acid metabolism may represent another pathway linking SSB to BC. Evidence suggests that increased basal hepatic fatty acids synthesis is one of the earliest metabolic changes induced by SSB consumption, preceding hypertriglyceridemia, hyperglycemia, or hyperinsulinemia [69]. Additionally, SSB consumption may alter PUFA metabolism through changes in acylcarnitine production and alterations in  $\beta$ -oxidation flux [70, 71]. Although limited evidence implicates PUFAs, particularly omega-6 fatty acids, in BC development [72, 73], the role of omega-6 fatty acids as potential mediators of SSB consumption and BC risk warrants further investigations.

### Strengths and limitations of this study

The major strength of our study was the use of two study designs, which yielded consistent findings. Moreover, this study provides the first evidence identifying anthropometric and metabolic mediators, such as BMI and uric acid, linking SSB consumption to BC risk. These findings highlight potential modifiable biomarkers that could inform prevention strategies for SSB-related BC. However, there are several limitations. First, beverage consumption was measured at a single time point, which might not fully capture long-term exposure. Nevertheless, the MR reflects lifelong differences in usual levels of exposure mitigating this concern [74]. Second, while



**Fig. 2** Mediation analysis on the causal association between sugar sweetened beverages and breast cancer using two-step Mendelian randomization. Abbreviations: CI, confidence interval; OR, odds ratio; LDL, low-density lipoprotein; HDL-C, high-density lipoprotein cholesterol; TFAs, total fatty acids. <sup>a</sup>The  $\beta$ /OR and 95% CI were estimated using two-sample univariable Mendelian randomization analysis, with inverse-variance weighted as the main method. \* $P < 0.05$ ; \*\* $P < 0.01$

extensive sensitivity analyses were performed, the possibility of pleiotropic effects cannot be entirely ruled out. For example, genetic variants associated with SSB consumption might also influence other dietary habits, introducing potential pleiotropy. Although observational analyses adjusted for multiple potential confounders, residual or unmeasured confounding could not be excluded. Therefore, further evidence is required to definitively establish causality. Third, our cohort study consisted exclusively of older Chinese women, which limits the generalizability of the findings to younger or more diverse populations. However, older women tend to have slower metabolism and reduced appetite [75], and SSB consumption is less prevalent in this demographic, potentially leading to an underestimation of the true association. Finally, the mediation analyses focused on widely recognized biomarkers, which may not fully capture all relevant pathways. Future studies should investigate additional potential mediators, to provide a more comprehensive understanding of the mechanisms underlying the observed associations.

**Conclusion**

Both observational and MR studies identified SSB consumption as a significant risk factor for BC, mediated by modifiable anthropo-metabolic markers such as BMI and uric acid. Our findings suggested a potential basis for targeted prevention strategies, implying that reducing SSB intake and addressing associated metabolic alternations may help mitigate BC risk.

Abbreviations	
SSB	Sugar sweetened beverages
PFJ	Pure fruit juice
BC	Breast cancer
MR	Mendelian randomization
GBCS	Guangzhou Biobank Cohort Study
GHHARE	Guangzhou Health and Happiness Association for the Respectable Elders
FFQ	Food Frequency Questionnaire
BMI	Body mass index
LDL-C	Low-density lipoprotein cholesterol
HDL-C	High-density lipoprotein cholesterol
HR	Hazard ratio
CI	Confidence interval
GWAS	Genome-wide association study
SNP	Single nucleotide polymorphism

IVW	Inverse-variance weighted
MR-PRESSO	MR pleiotropy residual sum and outlier
OR	Odds ratio
ER	Estrogen receptor
PUFA	Polyunsaturated fatty acid
TFA	Total fatty acid

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12937-025-01110-y>.

Additional file 1: Table S1. Beverage items used in the analysis. Table S2. Overview of genome-wide association study data. Table S3. Associations between beverage consumption and risk of breast cancer stratified by baseline age in the Guangzhou Biobank Cohort Study. Table S4. Associations between beverage consumption and risk of breast cancer stratified by baseline menopausal status in the Guangzhou Biobank Cohort Study. Table S5. Associations between beverage consumption and risk of breast cancer in 13,534 participants, excluding breast cancer or death events within the first year of follow-up in the Guangzhou Biobank Cohort Study. Table S6. Associations between beverage consumption and risk of breast cancer in the Guangzhou Biobank Cohort Study by proportional subdistribution hazards regression. Table S7. Associations between beverage consumption and risk of breast cancer in the Guangzhou Biobank Cohort Study using four partially adjusted models. Table S8. Associations between beverage consumption and 16 baseline anthropo-metabolic markers in the Guangzhou Biobank Cohort Study. Table S9. Mediation analysis with anthropo-metabolic markers as potential mediators for the association between beverage consumption and risk of breast cancer in the Guangzhou Biobank Cohort Study. Table S10. Summary information of the single nucleotide polymorphisms (SNPs) used as instrumental variables for beverage consumption in univariable Mendelian randomization. Table S11. Mendelian randomization estimates for the causal associations between sugar sweetened beverages consumption and estrogen receptor (ER)-positive/negative breast cancer. Table S12. Mendelian randomization estimates for the causal associations between sugar sweetened beverages consumption and anthropo-metabolic biomarkers. Table S13. Mendelian randomization estimates for the causal associations between anthropo-metabolic biomarkers and overall breast cancer. Figure S1. Flow diagram of participants selection.

## Acknowledgements

We gratefully acknowledge all the investigators and participants from the GBSC (Guangzhou Biobank Cohort Study).

## Authors' contributions

XYL, BHL, THL, KKC, WSZ and LX have substantial contributions to conception and design, acquisition of funding, data and interpretation of data; XYL and LX analyzed the data and drafted the article, BHL, THL, KKC and WSZ revised it critically for important intellectual content. All authors read and approved the final manuscript.

## Funding

This work was funded by the National Natural Science Foundation of China (82373661) and the Special Foundation for Science and Technology Basic Research Program (2019FY101103). The Guangzhou Biobank Cohort Study was funded by the University of Hong Kong Foundation for Educational Development and Research (SN/1f/HKUF-DC; C20400.28505200), the Health Medical Research Fund (HMRf/13143241) in Hong Kong, the Guangzhou Public Health Bureau (201102A211004011), the Natural Science Foundation of Guangdong (2018A030313140), Guangzhou Twelfth People's Hospital, the University of Birmingham, UK, School of Public Health, the University of Hong Kong and Greater Bay Area Public Health Research Collaboration. The funders of the study had no role in the study design; data collection, analysis, and interpretation; or writing of the report.

## Data availability

(1) Observational study: Data that support findings are not publicly available due to the protection of the privacy of participants, and are available from the Guangzhou Biobank Cohort Study Data Access Committee (gbcdata@hku.hk) on reasonable request. (2) Mendelian randomization study: The GWAS summary statistics are publicly available through OpenGWAS or GWAS catalog.

## Declarations

### Ethics approval and consent to participate

The Guangzhou Medical Ethics Committee of the Chinese Medical Association approved the study. All participants provided written, informed consent before participation. The study was performed in accordance with the Declaration of Helsinki.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>School of Public Health, Sun Yat-sen University, No. 74 Zhongshan 2nd Road, Guangzhou, Guangdong Province, China. <sup>2</sup>Guangzhou Center for Disease Control and Prevention, Guangzhou, China. <sup>3</sup>School of Public Health, the University of Hong Kong, Hong Kong, China. <sup>4</sup>Guangzhou Twelfth People's Hospital, Guangzhou, China. <sup>5</sup>School of Health Sciences, College of Medicine and Health, University of Birmingham, Birmingham, UK. <sup>6</sup>Greater Bay Area, Greater Bay Area Public Health Research Collaboration, Guangzhou, China.

Received: 21 January 2025 Accepted: 2 March 2025

Published online: 22 March 2025

## References

- Lara-Castor L, Micha R, Cudhea F, Miller V, Shi P, Zhang J, Sharib JR, Erndt-Marino J, Cash SB, Mozaffarian D. Sugar-sweetened beverage intakes among adults between 1990 and 2018 in 185 countries. *Nat Commun*. 2023;14(1):5957.
- Quadra GR, Paranaíba JR, Vilas-Boas J, Roland F, Amado AM, Barros N, Dias RJP, Cardoso SJ. A global trend of caffeine consumption over time and related-environmental impacts. *Environ Pollut*. 2020;256:113343.
- Grout L, Baker MG, French N, Hales S. A Review of Potential Public Health Impacts Associated With the Global Dairy Sector. *GeoHealth*. 2020;4(2):e2019GH000213.
- Dehghan M, Mente A, Rangarajan S, Sheridan P, Mohan V, Iqbal R, Gupta R, Lear S, Wentzel-Viljoen E, Avezum A, et al. Association of dairy intake with cardiovascular disease and mortality in 21 countries from five continents (PURE): a prospective cohort study. *Lancet*. 2018;392(10161):2288–97.
- Chazelas E, Srour B, Desmetz E, Kesse-Guyot E, Julia C, Deschamps V, Druet-Pecollo N, Galan P, Hercberg S, Latino-Martel P, et al. Sugary drink consumption and risk of cancer: results from NutriNet-Santé prospective cohort. *BMJ (Clin Res ed)*. 2019;366:l2408.
- Rodríguez-Ayala M, Donat-Vargas C, Moreno-Franco B, Mérida DM, Ramón Banegas J, Rodríguez-Artalejo F, Guallar-Castillón P. Association of a healthy beverage score with total mortality in the adult population of Spain: A nationwide cohort study. *PLoS Med*. 2024;21(1):e1004337.
- Jiang Y, Xu T, Dong W, Chu C, Zhou M. Study on the death and disease burden caused by high sugar-sweetened beverages intake in China from 1990 to 2019. *Eur J Public Health*. 2022;32(5):773–8.
- Singh GM, Micha R, Khatibzadeh S, Shi P, Lim S, Andrews KG, Engell RE, Ezzati M, Mozaffarian D. Global, Regional, and National Consumption of Sugar-Sweetened Beverages, Fruit Juices, and Milk: A Systematic Assessment of Beverage Intake in 187 Countries. *PLoS ONE*. 2015;10(8):e0124845.
- Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. 2024;74(1):12–49.

10. Lei S, Zheng R, Zhang S, Wang S, Chen R, Sun K, Zeng H, Zhou J, Wei W. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. *Cancer Commun* (London, England). 2021;41(11):1183–94.
11. Cao W, Chen HD, Yu YW, Li N, Chen WQ. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. *Chin Med J*. 2021;134(7):783–91.
12. Yin M, Wang F, Zhang Y, Meng R, Yuan X, Wang Q, Yu Y. Analysis on Incidence and Mortality Trends and Age-Period-Cohort of Breast Cancer in Chinese Women from 1990 to 2019. *Int J Environ Res Public Health*. 2023;20(1):826.
13. Pan B, Lai H, Ma N, Li D, Deng X, Wang X, Zhang Q, Yang Q, Wang Q, Zhu H, et al. Association of soft drinks and 100% fruit juice consumption with risk of cancer: a systematic review and dose-response meta-analysis of prospective cohort studies. *Int J Behav Nutr Physic Activity*. 2023;20(1):58.
14. Llahfa F, Gil-Lespinard M, Unal P, de Villasante I, Castañeda J, Zamora-Ros R. Consumption of Sweet Beverages and Cancer Risk. A Systematic Review and Meta-Analysis of Observational Studies. *Nutrients*. 2021;13(2):516.
15. Shin S, Fu J, Shin WK, Huang D, Min S, Kang D. Association of food groups and dietary pattern with breast cancer risk: A systematic review and meta-analysis. *Clin Nutr* (Edinburgh, Scotland). 2023;42(3):282–97.
16. Kazemi A, Barati-Boldaji R, Soltani S, Mohammadipoor N, Esmaeilinezhad Z, Clark CCT, Babajafari S, Akbarzadeh M. Intake of Various Food Groups and Risk of Breast Cancer: A Systematic Review and Dose-Response Meta-Analysis of Prospective Studies. *Adv Nutr* (Bethesda, Md). 2021;12(3):809–49.
17. Lumsden AL, Mulugeta A, Hyppönen E. Milk consumption and risk of twelve cancers: A large-scale observational and Mendelian randomisation study. *Clin Nutr* (Edinburgh, Scotland). 2023;42(1):1–8.
18. Gianfredi V, Nucci D, Abalsamo A, Acito M, Villarini M, Moretti M, Realdon S. Green Tea Consumption and Risk of Breast Cancer and Recurrence-A Systematic Review and Meta-Analysis of Observational Studies. *Nutrients*. 2018;10(12):1886.
19. Bhoo-Pathy N, Peeters PH, Uiterwaal CS, Bueno-de-Mesquita HB, Bulgiba AM, Bech BH, Overvad K, Tjønneland A, Olsen A, Clavel-Chapelon F, et al. Coffee and tea consumption and risk of pre- and postmenopausal breast cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort study. *Breast Cancer Res*. 2015;17(1):15.
20. Hodge AM, Bassett JK, Milne RL, English DR, Giles GG. Consumption of sugar-sweetened and artificially sweetened soft drinks and risk of obesity-related cancers. *Public Health Nutr*. 2018;21(9):1618–26.
21. Ma J, Jacques PF, Meigs JB, Fox CS, Rogers GT, Smith CE, Hruby A, Saltzman E, McKeown NM. Sugar-Sweetened Beverage but Not Diet Soda Consumption Is Positively Associated with Progression of Insulin Resistance and Prediabetes. *J Nutr*. 2016;146(12):2544–50.
22. Doyle SL, Donohoe CL, Lysaght J, Reynolds JV. Visceral obesity, metabolic syndrome, insulin resistance and cancer. *Proc Nutr Soc*. 2012;71(1):181–9.
23. Aboriza M, Amalia R, Zuhrotun A, Ikram NKK, Novitasari D, Mughtaridi M. Coffee Bean and Its Chemical Constituent Caffeine and Chlorogenic Acid as Promising Chemoprevention Agents: Updated Biological Studies against Cancer Cells. *Molecules* (Basel, Switzerland). 2024;29(14):3302.
24. Latronico T, Petraglia T, Sileo C, Bilancia D, Rossano R, Luzzi GM. Inhibition of MMP-2 and MMP-9 by Dietary Antioxidants in THP-1 Macrophages and Sera from Patients with Breast Cancer. *Molecules* (Basel, Switzerland). 2024;29(8):1718.
25. Smith GD, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1–22.
26. Carter AR, Sanderson E, Hammerton G, Richmond RC, Davey Smith G, Heron J, Taylor AE, Davies NM, Howe LD. Mendelian randomisation for mediation analysis: current methods and challenges for implementation. *Eur J Epidemiol*. 2021;36(5):465–78.
27. Jiang C, Thomas GN, Lam TH, Schooling CM, Zhang W, Lao X, Adab P, Liu B, Leung GM, Cheng KK. Cohort profile: The Guangzhou Biobank Cohort Study, a Guangzhou-Hong Kong-Birmingham collaboration. *Int J Epidemiol*. 2006;35(4):844–52.
28. Woo J, Leung SSF, Ho SC, Lam TH, Janus ED. A food frequency questionnaire for use in the Chinese population in Hong Kong: Description and examination of validity. *Nutr Res*. 1997;17(11):1633–41.
29. Park JW, Han K, Shin DW, Yeo Y, Chang JW, Yoo JE, Jeong SM, Lee SK, Ryu JM, Park YM. Obesity and breast cancer risk for pre- and postmenopausal women among over 6 million Korean women. *Breast Cancer Res Treat*. 2021;185(2):495–506.
30. Lécuyer L, Victor Bala A, Deschasaux M, Bouchemal N, Nawfal Triba M, Vasson MP, Rossary A, Demidem A, Galan P, Hercberg S, et al. NMR metabolomic signatures reveal predictive plasma metabolites associated with long-term risk of developing breast cancer. *Int J Epidemiol*. 2018;47(2):484–94.
31. Seyed Khoei N, Carreras-Torres R, Murphy N, Gunter MJ, Brennan P, Smith-Byrne K, Mariosa D, McKay J, O'Mara TA, Jarrett R, et al. Genetically Raised Circulating Bilirubin Levels and Risk of Ten Cancers: A Mendelian Randomization Study. *Cells*. 2021;10(2):394.
32. Liu X, Morelli D, Littlejohns TJ, Clifton DA, Clifton L. Combining machine learning with Cox models to identify predictors for incident post-menopausal breast cancer in the UK Biobank. *Sci Rep*. 2023;13(1):9221.
33. Feng Y, Fu M, Guan X, Wang C, Yuan F, Bai Y, Meng H, Li G, Wei W, Li H, et al. Uric Acid Mediated the Association Between BMI and Postmenopausal Breast Cancer Incidence: A Bidirectional Mendelian Randomization Analysis and Prospective Cohort Study. *Front Endocrinol* (Lausanne). 2021;12:742411.
34. Zhong VW, Kuang A, Danning RD, Kraft P, van Dam RM, Chasman DI, Cornelis MC. A genome-wide association study of bitter and sweet beverage consumption. *Hum Mol Genet*. 2019;28(14):2449–57.
35. Cole JB, Florez JC, Hirschhorn JN. Comprehensive genomic analysis of dietary habits in UK Biobank identifies hundreds of genetic associations. *Nat Commun*. 2020;11(1):1467.
36. Michailidou K, Lindström S, Dennis J, Beesley J, Hui S, Kar S, Lemaçon A, Soucy P, Glubb D, Rostamianfar A, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017;551(7678):92–4.
37. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, Powell C, Vedantam S, Buchkovich ML, Yang J, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518(7538):197–206.
38. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Mägi R, Strawbridge RJ, Pers TH, Fischer K, Justice AE, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*. 2015;518(7538):187–96.
39. Chen J, Spracklen CN, Marenne G, Varshney A, Corbin LJ, Luan J, Willems SM, Wu Y, Zhang X, Horikoshi M, et al. The trans-ancestral genomic architecture of glycemic traits. *Nat Genet*. 2021;53(6):840–60.
40. Karjalainen MK, Karthikeyan S, Oliver-Williams C, Sliz E, Allara E, Fung WT, Surendran P, Zhang W, Jousilahti P, Kristiansson K, et al. Genome-wide characterization of circulating metabolic biomarkers. *Nature*. 2024;628(8006):130–8.
41. Dennis JK, Sealock JM, Straub P, Lee YH, Hucks D, Actkins K, Faucon A, Feng YA, Ge T, Goleva SB, et al. Clinical laboratory test-wide association scan of polygenic scores identifies biomarkers of complex disease. *Genome Med*. 2021;13(1):6.
42. Chen Y, Lu T, Pettersson-Kymmer U, Stewart ID, Butler-Laporte G, Nakani-shi T, Cerani A, Liang KYH, Yoshiji S, Willett JDS, et al. Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases. *Nat Genet*. 2023;55(1):44–53.
43. Köttgen A, Albrecht E, Teumer A, Vitart V, Krumsiek J, Hundertmark C, Pistis G, Ruggiero D, O'Seaghdha CM, Haller T, et al. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat Genet*. 2013;45(2):145–54.
44. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol*. 2009;170(2):244–56.
45. Nevo D, Liao X, Spiegelman D. Estimation and Inference for the Mediation Proportion. *Int J Biostat*. 2017;13(2):/ijb.2017.13.issue-2/ijb-2017-0006/ijb-2017-0006.xml.
46. Lin DY, Fleming TR, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. *Stat Med*. 1997;16(13):1515–27.
47. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016;40(4):304–14.
48. Cohen JF, Chalumeau M, Cohen R, Korevaar DA, Khoshnood B, Bossuyt PM. Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. *J Clin Epidemiol*. 2015;68(3):299–306.
49. Zhang Z, Wang L. Methods for evaluating mediation effects: Rationale and comparison. *New Trends Psychometr*. 2008;595:604.

50. Romanos-Nanclares A, Toledo E, Gardeazabal I, Jiménez-Moleón JJ, Martínez-González MA, Gea A. Sugar-sweetened beverage consumption and incidence of breast cancer: the Seguimiento Universidad de Navarra (SUN) Project. *Eur J Nutr*. 2019;58(7):2875–86.
51. Romanos-Nanclares A, Collins LC, Hu FB, Willett WC, Rosner BA, Toledo E, Eliassen AH. Sugar-Sweetened Beverages, Artificially Sweetened Beverages, and Breast Cancer Risk: Results From 2 Prospective US Cohorts. *J Nutr*. 2021;151(9):2768–79.
52. Arthur RS, Kirsh VA, Mossavar-Rahmani Y, Xue X, Rohan TE. Sugar-containing beverages and their association with risk of breast, endometrial, ovarian and colorectal cancers among Canadian women. *Cancer Epidemiol*. 2021;70:101855.
53. Riseberg E, Wu Y, Lam WC, Eliassen AH, Wang M, Zhang X, Willett WC, Smith-Warner SA. Lifetime dairy product consumption and breast cancer risk: a prospective cohort study by tumor subtypes. *Am J Clin Nutr*. 2024;119(2):302–13.
54. Fraser GE, Jaceldo-Siegl K, Orlich M, Mashchak A, Sirirat R, Knutsen S. Dairy, soy, and risk of breast cancer: those confounded milks. *Int J Epidemiol*. 2020;49(5):1526–37.
55. Gapstur SM, Gaudet MM, Wang Y, Hodge RA, Um CY, Hartman TJ, McCullough ML. Coffee Consumption and Invasive Breast Cancer Incidence among Postmenopausal Women in the Cancer Prevention Study-II Nutrition Cohort. *Cancer Epidemiol Biomarkers Prev*. 2020;29(11):2383–6.
56. Iwase M, Matsuo K, Koyanagi YNY, Ito H, Tamakoshi A, Wang C, Utada M, Ozasa K, Sugawara Y, Tsuji I, et al. Alcohol consumption and breast cancer risk in Japan: A pooled analysis of eight population-based cohort studies. *Int J Cancer*. 2021;148(11):2736–47.
57. Chlebowski RT, Chen Z, Anderson GL, Rohan T, Aragaki A, Lane D, Dolan NC, Paskett ED, McTiernan A, Hubbell FA, et al. Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *J Natl Cancer Inst*. 2005;97(6):439–48.
58. Qi Q, Chu AY, Kang JH, Jensen MK, Curhan GC, Pasquale LR, Ridker PM, Hunter DJ, Willett WC, Rimm EB, et al. Sugar-sweetened beverages and genetic risk of obesity. *N Engl J Med*. 2012;367(15):1387–96.
59. Lane MM, Travica N, Gamage E, Marshall S, Trakman GL, Young C, Teasdale SB, Dissanayaka T, Dawson SL, Orr R, et al. Sugar-Sweetened Beverages and Adverse Human Health Outcomes: An Umbrella Review of Meta-Analyses of Observational Studies. *Annu Rev Nutr*. 2024;44(1):383–404.
60. Choi HK, Mount DB, Reginato AM. Pathogenesis of gout. *Ann Intern Med*. 2005;143(7):499–516.
61. Ayoub-Charette S, Liu Q, Khan TA, Au-Yeung F, Blanco Mejia S, de Souza RJ, Wolever TM, Leiter LA, Kendall C, Sievenpiper JL. Important food sources of fructose-containing sugars and incident gout: a systematic review and meta-analysis of prospective cohort studies. *BMJ Open*. 2019;9(5):e024171.
62. Hong C, Yan Y, Su L, Chen D, Zhang C. Development of a risk-stratification scoring system for predicting risk of breast cancer based on non-alcoholic fatty liver disease, non-alcoholic fatty pancreas disease, and uric acid. *Open medicine (Warsaw, Poland)*. 2022;17(1):619–25.
63. Xue X, Sun Z, Ji X, Lin H, Jing H, Yu Q. Associations between serum uric acid and breast cancer incidence: A systematic review and meta-analysis. *Am J Med Sci*. 2024;368(6):610–20.
64. Kurajoh M, Fukumoto S, Yoshida S, Akari S, Murase T, Nakamura T, Ishii H, Yoshida H, Nagata Y, Morioka T, et al. Uric acid shown to contribute to increased oxidative stress level independent of xanthine oxidoreductase activity in MedCity21 health examination registry. *Sci Rep*. 2021;11(1):7378.
65. Čipak Gašparović A, Milković L, Dandachi N, Stanzer S, Pezdirc I, Vrančić J, Šitić S, Suppan C, Balic M. Chronic Oxidative Stress Promotes Molecular Changes Associated with Epithelial Mesenchymal Transition, NRF2, and Breast Cancer Stem Cell Phenotype. *Antioxidants (Basel, Switzerland)*. 2019;8(12):633.
66. Dalbeth N, Phipps-Green A, House ME, Gamble GD, Horne A, Stamp LK, Merriman TR. Body mass index modulates the relationship of sugar-sweetened beverage intake with serum urate concentrations and gout. *Arthritis Res Ther*. 2015;17(1):263.
67. Beeghly-Fadiel A, Khankari NK, Delahanty RJ, Shu XO, Lu Y, Schmidt MK, Bolla MK, Michailidou K, Wang Q, Dennis J, et al. A Mendelian randomization analysis of circulating lipid traits and breast cancer risk. *Int J Epidemiol*. 2020;49(4):1117–31.
68. Haslam DE, Chasman DI, Peloso GM, Herman MA, Dupuis J, Lichtenstein AH, Smith CE, Ridker PM, Jacques PF, Mora S, et al. Sugar-Sweetened Beverage Consumption and Plasma Lipoprotein Cholesterol, Apolipoprotein, and Lipoprotein Particle Size Concentrations in US Adults. *J Nutr*. 2022;152(11):2534–45.
69. Geidl-Flueck B, Hochuli M, Németh Á, Eberl A, Derron N, Köfeler HC, Tappy L, Berneis K, Spinass GA, Gerber PA. Fructose- and sucrose- but not glucose-sweetened beverages promote hepatic de novo lipogenesis: A randomized controlled trial. *J Hepatol*. 2021;75(1):46–54.
70. Hochuli M, Aeberli I, Weiss A, Hersberger M, Troxler H, Gerber PA, Spinass GA, Berneis K. Sugar-sweetened beverages with moderate amounts of fructose, but not sucrose, induce Fatty Acid synthesis in healthy young men: a randomized crossover study. *J Clin Endocrinol Metab*. 2014;99(6):2164–72.
71. Dambrova M, Makrecka-Kuka M, Kuka J, Vilskersts R, Nordberg D, Attwood MM, Smesny S, Sen ZD, Guo AC, Oler E, et al. Acylcarnitines: Nomenclature, Biomarkers, Therapeutic Potential, Drug Targets, and Clinical Trials. *Pharmacol Rev*. 2022;74(3):506–51.
72. Pouchieu C, Chajès V, Laporte F, Kesse-Guyot E, Galan P, Hercberg S, Latino-Martel P, Touvier M. Prospective associations between plasma saturated, monounsaturated and polyunsaturated fatty acids and overall and breast cancer risk - modulation by antioxidants: a nested case-control study. *PLoS ONE*. 2014;9(2):e90442.
73. Simonsen N, van't Veer P, Strain JJ, Martin-Moreno JM, Huttunen JK, Navajas JF, Martin BC, Thamm M, Kardinaal AF, Kok FJ, et al. Adipose tissue omega-3 and omega-6 fatty acid content and breast cancer in the EURAMIC study. European Community Multicenter Study on Antioxidants, Myocardial Infarction, and Breast Cancer. *Am J Epidemiol*. 1998;147(4):342–52.
74. Nazarzadeh M, Pinho-Gomes AC, Bidet Z, Dehghan A, Canoy D, Hassaine A, Ayala Solares JR, Salimi-Khorshidi G, Smith GD, Otto CM, et al. Plasma lipids and risk of aortic valve stenosis: a Mendelian randomization study. *Eur Heart J*. 2020;41(40):3913–20.
75. Zhao H, Andreyeva T. Diet Quality and Health in Older Americans. *Nutrients*. 2022;14(6):1198.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.