

Original Research Article

Healthy dietary patterns, genetic risk, and gastrointestinal cancer incident risk: a large-scale prospective cohort study



Yimin Cai^{1,†}, Canlin Hong^{2,†}, Jinxin Han^{3,†}, Linyun Fan^{2,†}, Xinyu Xiao^{2,†}, Jun Xiao^{4,†},
Yongchang Wei⁵, Ying Zhu², Jianbo Tian^{2,*}, Xu Zhu^{6,*}, Meng Jin^{7,*}, Xiaoping Miao^{1,2,*}

¹ Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China; ² Department of Epidemiology and Biostatistics, School of Public Health, Wuhan University, Wuhan, Hubei, China;

³ Department of Gastrointestinal Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China; ⁴ Department of Thyroid and Breast Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China; ⁵ Department of Gastrointestinal Oncology, Hubei Cancer Clinical Study Center, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, China; ⁶ Department of Gastrointestinal Surgery, Renmin Hospital, Wuhan University, Wuhan, Hubei, China; ⁷ Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

ABSTRACT

Background: Dietary patterns have been associated with several cancers, especially gastrointestinal cancer (GIC). However, whether a healthy dietary pattern could modify the risk of GIC among people with different genetic backgrounds is not clear.

Objective: The objective of the study was to investigate how dietary patterns and genetic susceptibility contribute to the risk of GIC independently and jointly.

Methods: This large-scale prospective cohort study included 105,463 participants in UK Biobank who were aged 40–72 y and cancer-free at baseline. Dietary intake (Oxford WebQ) was used to calculate dietary pattern scores including dietary approach to stop hypertension (DASH) score and healthful plant-based diet index (hPDI). Genetic risk was quantified by a polygenic risk score (PRS) comprising 129 known GIC-associated loci. Cox proportional hazards regression was performed to estimate the associations of dietary patterns and PRS with GIC incidence after adjusting for potential confounders.

Results: Over a median follow-up of 11.70 y, 1,661 participants were diagnosed with GIC. DASH and hPDI were associated with 20% and 36% reductions, respectively, in GIC risk. Low PRS was associated with a 30 % decrease in GIC risk (HR: 0.70; 95% CI: 0.62, 0.79). Participants with healthy dietary scores at high-genetic risk had a lower GIC risk with HR of 0.77 (95% CI: 0.60, 0.98) for DASH and 0.66 (95% CI: 0.52, 0.84) for hPDI than those with unhealthy dietary score. Participants with both high-dietary score and low-genetic risk showed the lowest risk of GIC, with HR of 0.58 (95% CI: 0.45, 0.75) for DASH and 0.45 (95% CI: 0.34, 0.58) for hPDI.

Conclusions: Adherence to DASH and hPDI were associated with a lower risk of some gastrointestinal cancers, and these 2 dietary patterns may partly compensate for genetic predispositions to cancer. Our results advance the development of precision medicine strategies that consider both dietary patterns and genetics to improve gastrointestinal health.

Keywords: gastrointestinal cancer, dietary patterns, genetic risk, gene-diet interaction

Introduction

Gastrointestinal cancer (GIC) is a major cause of global cancer incidence and mortality, accounting for approximately one-fourth of all cancer cases and one-third of all cancer deaths worldwide [1]. Diet is

believed to be an important modifiable factor associated with cancer risk. It is often evaluated by dietary patterns, which take multiple foods and dietary components together into consideration, providing a more comprehensive snapshot of dietary exposures. Accumulating evidence suggests that healthy dietary patterns can reduce the risk of site-specific GICs, such as esophageal cancer (ESC) [2], gastric cancer (GC) [2,3],

Abbreviations: CI, confidence interval; CRC, colorectal cancer; DASH, dietary approach to stop hypertension; ESC, esophageal cancer; GC, gastric cancer; GIC, gastrointestinal cancer; GWASs, genome-wide association studies; hPDI, healthful plant-based diet index; HR, hazard ratio; MSI, microsatellite instable; PAC, pancreatic cancer; PRS, polygenic risk score; SNPs, single nucleotide polymorphisms; TNM, tumor node metastasis.

* Corresponding author.

E-mail addresses: tianjb@whu.edu.cn (J. Tian), zhuxuwhu@whu.edu.cn (X. Zhu), Mengjin@tjh.tjmu.edu.cn (M. Jin), miaoxp@hust.edu.cn (X. Miao).

† YC, CH, JH, LF, XX, and JX contributed equally to this work.

<https://doi.org/10.1016/j.ajcnut.2023.11.015>

Received 14 July 2023; Received in revised form 19 November 2023; Accepted 28 November 2023; Available online 30 November 2023
0002-9165/© 2023 American Society for Nutrition. Published by Elsevier Inc. All rights reserved.

pancreatic cancer (PAC) [4], and colorectal cancer (CRC) [5,6]. Although many kinds of healthy dietary patterns had been associated with a lower risk of GIC in previous studies [7,8], limited long-term prospective studies have systematically examined whether the dietary approaches to stop hypertension (DASH) score and the healthful plant-based diet index (hPDI) render similar associations with GIC risk. DASH and hPDI represent 2 comprehensive dietary patterns with different focuses. DASH has been developed to target chronic disease prevention guidelines, whereas hPDI is focused on cultural ways of eating that are believed to be healthy. Investigating the effects of different dietary patterns on GIC in large prospective population-based cohorts is necessary to increase the understanding of food and nutrient combinations that may influence the development of GIC, which will also further the standardization of dietary pattern methodologies for assessing gastrointestinal health outcomes.

As is known to all, genetic predisposition can be identified at an early age and can contribute to a persistent effect on gastrointestinal health throughout adult life [9]. So far, a large number of susceptibility loci associated with site-specific GIC risk have been discovered by

genome-wide association studies (GWASs) [10,11]. Clumping these susceptibility loci with modest effect, polygenic risk scores (PRSs) can be constructed to effectively estimate the genetic risk of GIC. However, the predictive performance of PRS differs among individuals depending on their environmental exposure [10,11], suggesting that the genetic risk of GIC can be offset by nonheritable risk factors. It is vital to identify those environmental factors, particularly dietary patterns, which might modify genetic associations related to GIC [12]. Elucidating the gene-diet interactions could help us develop personalized nutrition guidance for effective prevention of GIC, especially among those with a strong genetic preposition. Thus, it is warranted to explore the joint effect and interaction of genetic risk and dietary patterns on GIC risk, which has been ignored by previous studies.

To illustrate the issues above, we first examined the associations of 2 dietary patterns (DASH and hPDI) with GIC. Subsequently, we constructed a PRS for GIC derived from individual PRSs of ESC, GC, PAC, and CRC to investigate the effect of genetic background on GIC risk. Furthermore, our study explored how dietary patterns and genetic factors jointly act in the development of GIC and the extent to which

TABLE 1

Baseline characteristics of 105,463 participants from UK Biobank

Characteristics	All participants	Participants without GIC	Participants with GIC	<i>P</i> value
Age (y), mean±SD	56.56 ± 7.83	56.49 ± 7.83	61.00 ± 6.24	<0.001
Gender				<0.001
Male	50,728 (48.10)	49,689 (47.87)	1,039 (62.55)	
Female	54,735 (51.90)	54,113 (52.13)	622 (37.45)	
Ethnicity				0.001
White	101,206 (95.96)	99,581 (95.93)	1,625 (97.84)	
Mixed	571 (0.54)	561 (0.54)	10 (0.60)	
Asian or Asian British	1,496 (1.42)	1,489 (1.43)	7 (0.42)	
Black or Black British	1,201 (1.14)	1,192 (1.15)	9 (0.54)	
Chinese	259 (0.25)	256 (0.25)	3 (0.18)	
Other	730 (0.69)	723 (0.70)	7 (0.42)	
Townsend deprivation index				0.077
Level 1 (least deprived)	21,083 (19.99)	20,721 (19.96)	362 (21.79)	
Level 2	21,102 (20.01)	20,759 (20.00)	343 (20.65)	
Level 3	21,093 (20.00)	20,753 (19.99)	340 (20.47)	
Level 4	21,094 (20.00)	20,802 (20.04)	292 (17.58)	
Level 5 (most deprived)	21,091 (20.00)	20,767 (20.01)	324 (19.51)	
Education level				<0.001
College or university degree	46,017 (43.63)	45,343 (43.68)	674 (40.58)	
A levels/AS levels or equivalent	13,826 (13.11)	13,643 (13.14)	183 (11.02)	
O levels/GCSEs or equivalent	21,519 (20.40)	21,161 (20.39)	358 (21.55)	
CSEs or equivalent	4,122 (3.92)	4,087 (3.94)	35 (2.11)	
NVQ or HND or HNC or equivalent	5,865 (5.56)	5,753 (5.54)	112 (6.74)	
Other	14,114 (13.38)	13,815 (13.31)	299 (18.00)	
Smoke				<0.001
Never	58,597 (55.56)	57,848 (55.73)	749 (45.09)	
Previous	39,164 (37.14)	38,397 (36.99)	767 (46.18)	
Current	7,702 (7.30)	7,557 (7.28)	145 (8.73)	
Alcohol				<0.001
Never	6,893 (6.53)	6,777 (6.53)	116 (6.98)	
Moderate	73,811 (69.99)	72,742 (70.08)	1,069 (64.36)	
Heavy	24,759 (23.48)	24,283 (23.39)	476 (28.66)	
BMI, mean±SD	27.17 ± 4.68	26.48 ± 4.68	27.96 ± 4.76	<0.001
MET-min/wk, mean±SD	2,439.02 ± 2,426.01	2,438.14 ± 2,422.82	2,493.75 ± 2,618.13	0.354
Energy intake, mean±SD	2,124.60 ± 699.80	2,123.88 ± 699.63	2,169.34 ± 709.10	0.009
Family history of GIC	11,762 (11.15)	11,522 (11.10)	240 (14.45)	<0.001
History of hypertension	33,477 (31.74)	32,771 (31.57)	706 (42.50)	<0.001
History of heart diseases	5,008 (4.75)	4,903 (4.72)	105 (6.32)	0.002
History of diabetes	5,547 (5.26)	5,384 (5.19)	163 (9.81)	<0.001

Abbreviations: GIC, gastrointestinal cancer; A levels/AS levels, advanced levels/advanced subsidiary levels; O levels/GCSE, ordinary levels/general certificate of secondary education; CSE, certificate of secondary education; NVQ, national vocational qualification; HND, higher national diploma; HNC, higher national certificate; MET, metabolic equivalent.

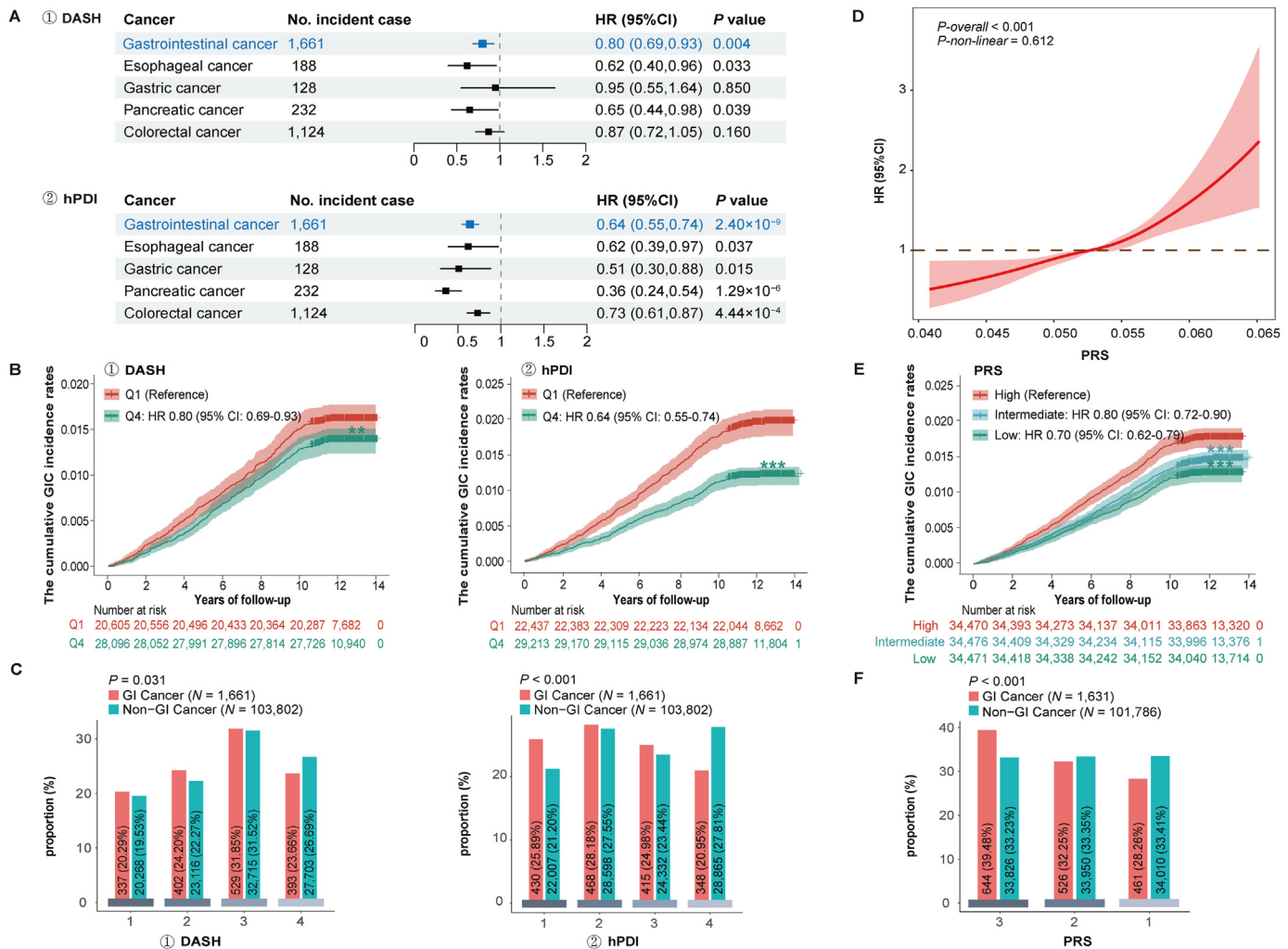


FIGURE 1. Effect of dietary patterns and genetic factors on the risk of GIC incidence. (A) The associations of dietary patterns with risk of GIC, ESC, GC, PAC, and CRC. Forest plots showed HRs for the highest category compared with the lowest category. (B) Cumulative GIC incidences for Quartile 1 and Quartile 4 of DASH score and hPDI. (C) Proportions of GI cancer and non-GI cancer were discriminated by quartiles of DASH score and hPDI. (D) The relationship of genetic factors with risk of GIC. A linear relationship between PRS and GIC risk was assessed using a restricted cubic spline analysis. (E) Cumulative GIC incidences for high, intermediate, and low-genetic risk groups. (F) Proportions of GI cancer and non-GI cancer were discriminated by tertiles of PRS. All HRs and 95% CIs were estimated using Cox proportional hazard models with adjustment for age, gender, ethnicity, Townsend deprivation index, education level, smoking status, alcohol consumption, BMI, physical activity, energy intake, family history of GIC, history of hypertension, heart diseases, and diabetes. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. CI, confidence interval; CRC, colorectal cancer; DASH, dietary approaches to stop hypertension; ESC, esophageal cancer; GC, gastric cancer; GIC, gastrointestinal cancer; hPDI, healthful plant-based diet index; HR, hazard ratio; PAC, pancreatic cancer; PRS, polygenic risk score.

GIC risk of individuals with high-genetic risk can be counteracted by healthy dietary patterns. Overall, our findings provide important evidence to support balancing dietary intake to an individual's genetic risk for GIC prevention and substantially advance public policies for improving dietary pattern to reduce the burden of GIC in the general population.

Methods

Study population

UK Biobank is a large-scale cohort study that recruited >500,000 adults from United Kingdom during 2007–2010 [13]. Details of the study design and methods have been published on the website of the UK Biobank [14]. Before analysis, we excluded participants who had missing information on diet ($N = 291,461$), other covariates ($N = 84$,

074), and follow-up ($N = 357$), who had a diagnosis of cancer at baseline ($N = 11,602$), who were pregnant ($N = 30$), who had energy intake <500 kcal or >6000 kcal ($N = 450$). Participants who often changed their diet ($N = 9,024$) were also excluded if they reported that their diets often varied much from week to week. Finally, a total of 105,463 participants remained in our analysis (Supplementary Figure 1).

Ascertainment of outcome

In our study, the primary outcome is the incident GIC, including cancers from the esophagus, stomach, pancreas, to colorectum. Diagnosis of GIC was ascertained according to the 10th Revision of the International Classification of Disease, including malignant neoplasm of the esophagus (C15), stomach (C16), colorectum (C18–20, exp C18.1), and pancreas (C25). The end point was defined as the date of

TABLE 2
The associations of dietary patterns and genetic risk with the risk of gastrointestinal cancer

Factors	Categories	No. of cases (person-years)	Model 1 ^d		P-trend ^e		Model 2 ^d		P-trend ^e		Model 3 ^d		P-trend ^e	
			HR (95% CI)	P value			HR (95% CI)	P value			HR (95% CI)	P value		
DASH ^a	1	337/242,101	1.00 (Ref)		0.008	1.00 (Ref)			2.57 × 10 ⁻⁵	1.00 (Ref)			0.002	
	2	402/276,829	1.04 (0.90, 1.21)	0.555		0.93 (0.80, 1.07)	0.321			0.95 (0.82, 1.10)	0.524			
	3	529/391,791	0.97 (0.85, 1.11)	0.681		0.84 (0.73, 0.96)	0.011			0.87 (0.76, 1.01)	0.061			
	4	393/331,287	0.85 (0.74, 0.99)	0.033		0.75 (0.64, 0.86)	9.35 × 10 ⁻⁵			0.80 (0.69, 0.93)	0.004			
hPDI ^b	Per 10-unit increment		0.86 (0.77, 0.96)	0.005	2.15 × 10 ⁻¹⁰	0.77 (0.69, 0.86)	4.33 × 10 ⁻⁶		3.73 × 10 ⁻¹⁰	0.82 (0.73, 0.91)	4.91 × 10 ⁻⁴		2.41 × 10 ⁻⁸	
	1	430/263,818	1.00 (Ref)			1.00 (Ref)				1.00 (Ref)				
	2	468/341,812	0.84 (0.74, 0.96)	0.009		0.80 (0.71, 0.92)	0.001			0.82 (0.72, 0.93)	0.003			
	3	415/291,015	0.87 (0.76, 1.00)	0.051		0.84 (0.73, 0.96)	0.010			0.86 (0.75, 0.98)	0.028			
Genetic risk ^c	4	348/345,363	0.62 (0.54, 0.71)	2.87 × 10 ⁻¹¹	2.24 × 10 ⁻⁸	0.62 (0.53, 0.71)	3.22 × 10 ⁻¹¹		7.12 × 10 ⁻⁹	0.64 (0.55, 0.74)	2.40 × 10 ⁻⁹		6.49 × 10 ⁻⁹	
	Per 10-unit increment		0.77 (0.71, 0.82)	1.05 × 10 ⁻¹²		0.76 (0.70, 0.82)	6.09 × 10 ⁻¹³			0.77 (0.71, 0.83)	8.48 × 10 ⁻¹¹			
	High	644/405,380	1.00 (Ref)			1.00 (Ref)				1.00 (Ref)				
	Intermediate	526/406,197	0.81 (0.73, 0.91)	4.98 × 10 ⁻⁴		0.80 (0.72, 0.90)	2.07 × 10 ⁻⁴			0.80 (0.72, 0.90)	2.15 × 10 ⁻⁴			
	Low	461/406,926	0.71 (0.63, 0.80)	3.13 × 10 ⁻⁸		0.71 (0.63, 0.79)	1.05 × 10 ⁻⁸			0.70 (0.62, 0.79)	9.43 × 10 ⁻⁹			

Abbreviations: CI, confidence interval; DASH, dietary approaches to stop hypertension; hPDI, healthful plant-based diet index; HR, hazard ratio; PRS, polygenic risk score; Ref, reference.

^a Quartiles of DASH scores: 1 (<17); 2 (17 to <20); 3 (20 to <24); 4 (≥24);

^b Quartiles of hPDI: 1 (<53); 2 (53 to <58); 3 (58 to <62); 4 (≥62);

^c Tertiles of PRS: high (≥0.054); intermediate (0.051 to <0.054); low (<0.051).

^d Model 1 was not adjusted. Model 2 was adjusted for age and gender. Model 3 was adjusted for age, gender, ethnicity, Townsend deprivation index, education level, smoking status, alcohol consumption, BMI, physical activity, energy intake, family history of GIC, history of hypertension, heart diseases, and diabetes.

^e Test for trend based on variable containing median value for each quartile (DASH, hPDI) or tertile (PRS).

first-ever any GIC diagnosis through linkage with national cancer registries, or the end of follow-up on 31 March 2021.

Assessment of dietary patterns

Dietary data were collected through the 24-h dietary recall (Oxford WebQ) [15]. We calculated the quantity of each food or drink consumed by multiplying the assigned portion size of each food or beverage by the amount consumed [16]. Intakes of energy and nutrients were estimated from those foods and drinks according to the amount consumed and the energy and nutrient composition [17]. UK Biobank collected dietary data at baseline and on 4 separate occasions at regular intervals between February 2011 and June 2012 [18]. For each participant, we used one (rather than multiple) dietary recall at the baseline assessment to calculate the scores of 2 popular dietary patterns: DASH and hPDI. Alternatively, the utilization of multiple dietary recalls might have introduced a bias known as reverse causality, where at least one dietary exposure point would have proceeded with the evaluation of outcome measures [19].

The components and scoring criteria for DASH and hPDI are summarized in detail in [Supplementary Tables 1 and 2](#). DASH consists of 8 components designed by Fung et al. [20]. hPDI was scored according to 18 components, which is similar to previous studies [21]. Of note, some modifications were made to accommodate the calculation of the dietary information in the study. For the DASH score, the intake of sodium was not collected in UK Biobank, and therefore, this component was not calculated in our study. For the hPDI, data on the intake of vegetable oil were unavailable in UK Biobank, therefore, we adapted the hPDI to score the intake of olive oil instead. The intake of each component in DASH and hPDI was divided into quintiles and given positive or reverse scores. For 5 components (fruits, vegetables, nuts and legumes, whole grains, and low-fat dairy products) in DASH, participants received 1 point for the lowest quintile and 5 points for the highest quintile. Conversely, for the other 2 components (red and processed meats, sweetened beverages), reverse scoring for quintiles was used. For hPDI, 7 components (vegetable oils, whole grains, fruits, vegetables, nuts, legumes and vegetarian protein alternatives, and tea and coffee) were calculated with positive scoring that a score of one was assigned to participants who had the lowest intake and a score of 5 was assigned to participants who had the highest intake. Other components were calculated with reverse scoring. The total DASH score and hPDI ranged from 7 to 35 and 18 to 90, respectively, with a higher score reflecting higher adherence to DASH and hPDI.

Construction of PRS

To evaluate the genetic predisposition of GIC, we constructed a PRS based on 129 single nucleotide polymorphisms (SNPs) that reached the genome-wide significance threshold ($P < 5 \times 10^{-8}$) in populations of European ancestry. All the included SNPs were extracted from the primary analyses [10,11,22,23]. Detailed information on the SNPs is provided in [Supplementary Table 3](#). The calculation of PRS was performed by using PRSice-2 algorithm [24] with the following formula:

$$PRS_j = \sum_{i=1}^n \beta_i G_{ij}$$

Where β value (ln(OR)—natural logarithm transformed odds ratio) is the summary statistic for the effective allele and G is the number of the effective allele observed. The calculated PRS was stratified into low (the lowest tertile of PRS), intermediate (the second tertile), and high (the highest tertile) risk groups.

TABLE 3
The associations between dietary patterns and the risk of individual gastrointestinal cancers

Factors	Outcomes	Model 1 ^a		Model 2 ^a		Model 3 ^a	
		HR (95% CI) ^b	<i>P</i> value	HR (95% CI) ^b	<i>P</i> value	HR (95% CI) ^b	<i>P</i> value
DASH	Esophageal cancer	0.52 (0.34, 0.79)	0.002	0.51 (0.33, 0.78)	0.002	0.62 (0.40, 0.96)	0.033
	Gastric cancer	0.94 (0.56, 1.58)	0.809	0.85 (0.50, 1.44)	0.546	0.95 (0.55, 1.64)	0.850
	Pancreatic cancer	0.72 (0.49, 1.06)	0.094	0.56 (0.38, 0.83)	0.004	0.65 (0.44, 0.98)	0.039
	Colorectal cancer	0.97 (0.81, 1.16)	0.731	0.85 (0.71, 1.02)	0.077	0.87 (0.72, 1.05)	0.160
hPDI	Esophageal cancer	0.47 (0.30, 0.73)	0.001	0.55 (0.35, 0.86)	0.009	0.62 (0.39, 0.97)	0.037
	Gastric cancer	0.42 (0.25, 0.71)	0.001	0.43 (0.25, 0.73)	0.002	0.51 (0.30, 0.88)	0.015
	Pancreatic cancer	0.42 (0.28, 0.63)	2.74×10^{-5}	0.37 (0.25, 0.56)	2.30×10^{-6}	0.36 (0.24, 0.54)	1.29×10^{-6}
	Colorectal cancer	0.72 (0.60, 0.85)	1.19×10^{-4}	0.71 (0.60, 0.84)	9.99×10^{-5}	0.73 (0.61, 0.87)	4.44×10^{-4}

Abbreviations: CI, confidence interval; DASH, dietary approaches to stop hypertension; GIC, gastrointestinal cancer; hPDI, healthful plant-based diet index; HR, hazard ratio.

^a Model 1 was not adjusted. Model 2 was adjusted for age and gender. Model 3 was adjusted for age, gender, ethnicity, Townsend deprivation index, education level, smoking status, alcohol consumption, BMI, physical activity, energy intake, family history of GIC, history of hypertension, heart diseases, and diabetes.

^b HRs were shown for the highest category of dietary score compared with the lowest category.

Covariates

The definitions of all included covariates are listed in [Supplementary Table 4](#). The covariates, which could confuse the effect of dietary intake on GIC potentially, were selected based on prior knowledge and previous studies [2,3,25]. At enrollment, baseline touchscreen questionnaires collected large amounts of participants' information on sociodemographic characteristics, lifestyle exposures, and medical history, including age, gender, ethnicity, education level, smoking status, alcohol consumption, physical activity, energy intake, family history of GIC, history of hypertension, heart diseases, and diabetes. Participants were asked to select their ethnicity from options of "White," "Mixed," "Asian or Asian British," "Black or Black British," "Chinese," "other ethnic group," "do not know," or "prefer not to answer." The mixed category combined the responses of the UK Biobank ethnicity questions of "mixed," "White and Black Caribbean," "White and Black African," "White and Asian," and "any other mixed background." Education was assessed by asking "Which of the following qualifications do you have?," with the options of "college or university degree," "A levels/AS levels or equivalent," "O levels/GCSEs or equivalent," "CSEs or equivalent," "NVQ/HND/HNC or equivalent," "other professional qualifications (e.g., nursing or teaching)." Smoking status was self-reported as never, previous, and current smoker. Alcohol consumption was assessed through the question "About how often do you drink alcohol?," and categorized as never, moderate (3 or 4 times a week, once or twice a week, 1–3 times a month, special occasions only), and heavy (daily or almost daily). According to an adapted version of the short International Physical Activity Questionnaire [26], physical activity was assessed using the metabolic equivalent minutes per week, which were evaluated by the time spent in walking and moderate and vigorous activities. Townsend deprivation index, created by the UK Biobank group, is a composite measure of deprivation derived from national census data about unemployment, noncar ownership, nonhome ownership, and household overcrowding [27]. We operationalized the Townsend deprivation index as quintiles and a negative value of the Townsend deprivation index represented high socioeconomic status. Besides, BMI was calculated as weight/height² (kg/m²) measured at the assessment center.

Statistical analysis

Descriptive analyses of participant characteristics included mean and SD for continuous variables, and number (percentage) for categorical variables. Between-group differences in baseline characteristics were assessed using a *t*-test or Analysis of Variance test for continuous variables, and a Chi-square test for categorical variables, as appropriate. Cox proportional hazards regression was used to evaluate the associations of dietary patterns and PRS with GIC incidence by hazard ratios (HRs) and 95% confidence intervals (CIs). The assumption of proportional hazards was confirmed using Schoenfeld residual methods. Three types of models were built to adjust for potential confounders, which were added in a stepwise manner. Model 1 was without adjustment. Model 2 was adjusted for age and gender. In model 3, we further adjusted for ethnicity, Townsend deprivation index, education level, smoking status, alcohol consumption, BMI, physical activity, energy intake, family history of GIC, history of hypertension, heart diseases, and diabetes. To estimate the *P*-trend, we assigned the median value to the participants within the category. Kaplan–Meier analysis with log-rank testing was performed to show the cumulative incidence rate across the different categories of dietary patterns or PRS in total participants. Additionally, we examined the potential nonlinear relation between PRS and GIC risk with restricted cubic splines.

In subgroup analyses, the models were reanalyzed stratified by age, gender, education level, smoking status, alcohol consumption, obese, and family history. To address the possibility of reverse causality, we conducted lag analyses by applying different lag periods between dietary patterns and GIC incident risk (2-, 4-, 6-, and 8-y lag). We also performed sensitivity analysis to examine the robustness of our findings by excluding participants with BMI below the 1st or above the 99th centiles. To investigate whether the associations of dietary patterns with GIC might be different according to genetic risk, we cross-classified participants in accordance with categories of dietary patterns (quartile 1, and quartile 4) and PRS (high, intermediate, and low) and conducted joint analyses. We subsequently added a cross-product term into the model to test for multiplicative interaction between dietary patterns and genetic risk. The 95% CIs for the cumulative risks were derived by drawing 1000 bootstrap samples from the estimation data set. In addition, the integration values of dietary patterns and PRS were derived by obtaining predictions from the fitted generalized linear model object.

All statistical analyses were performed using R software (version 4.2.1). Statistical tests were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Characteristics of the included participants

Our study included 105,463 participants (mean age 56.56 y, 48.10 % males). During a median follow-up of 11.70 y, 1,661 GIC cases were diagnosed. Participants with GIC were composed of more males (62.55%), current smokers (8.73%), heavy drinkers (28.66%), and more participants who had hypertension (42.50%), heart diseases (6.32%), and diabetes history (9.81%) than participants without GIC (Table 1). We also compared the baseline characteristics of participants according to the quartiles of dietary score. As shown in Supplementary Table 5, participants with higher DASH scores were more likely to be older (mean age 57.60 y), females (61.26%), higher-educated (51.64% with college or university degree), never smokers (58.70%), and never drinkers (6.93%). They were also less likely to have a history of hypertension (29.13%), heart disease (4.16%), and diabetes (4.04%). We found similar results when comparing the characteristics of participants according to the quartiles of hPDI. The detailed characteristics are presented in Supplementary Table 6.

Associations of dietary patterns with GIC risk

We first investigated the associations of dietary patterns with the risk of GIC and site-specific GICs including ESC, GC, PAC, and CRC. We found inverse associations between DASH score (HR_{Q4} compared with Q₁: 0.80; 95% CI: 0.69, 0.93; $P = 0.004$), hPDI (HR_{Q4} compared with Q₁: 0.64; 95% CI: 0.55, 0.74; $P = 2.40 \times 10^{-9}$) and GIC risk (Figure 1A, Table 2). Specifically, DASH score was only inversely associated with the risks of ESC (HR_{Q4} compared with Q₁: 0.62; 95% CI: 0.40, 0.96; $P = 0.033$) and PAC (HR_{Q4} compared with Q₁: 0.65; 95% CI: 0.44, 0.98; $P = 0.039$), whereas the highest quartile of hPDI was significantly associated with decreased risks of all 4 site-specific cancers, reducing the risks of ESC, GC, PAC, and CRC by 38%, 49%, 64%, and 27%, respectively (Figure 1A, Table 3). The results of multivariable-adjusted analyses for the associations between dietary patterns and GIC risk are presented in Table 2. In the crude model without adjustment, participants at quartile 4 of DASH score and hPDI had a significantly lower risk of GIC than those at quartile 1, with HRs of 0.85 (95% CI: 0.74, 0.99; $P = 0.033$) and 0.62 (95% CI: 0.54, 0.71; $P = 2.87 \times 10^{-11}$), respectively. The results remained stable and statistically significant after adjusting age and gender in model 2 (DASH: P -trend = 2.57×10^{-5} , hPDI: P -trend = 3.73×10^{-10}) and all confounding factors in model 3 (DASH: P -trend = 0.002, hPDI: P -trend = 2.41×10^{-8}). To further validate the inverse relationship over follow-up time, we conducted cumulative risk analyses, which showed that the cumulative incidence was higher for participants with low-dietary scores than those with high-dietary scores (Figure 1B). Besides, the distribution of dietary scores in GIC population differs from that in the non-GIC population (DASH: $P = 0.031$, hPDI: $P < 0.001$). Non-GIC populations were composed of a higher proportion of individuals with the highest quartile of dietary score (DASH: 26.69%, hPDI: 27.81%) (Figure 1C). Overall, these results demonstrated that greater adherence to healthy dietary patterns, such as DASH or hPDI may contribute to the decreased risk of GIC.

As shown in Supplementary Figure 2A, the effects of DASH on GIC risk were more significant among participants aged ≥ 60 y, male,

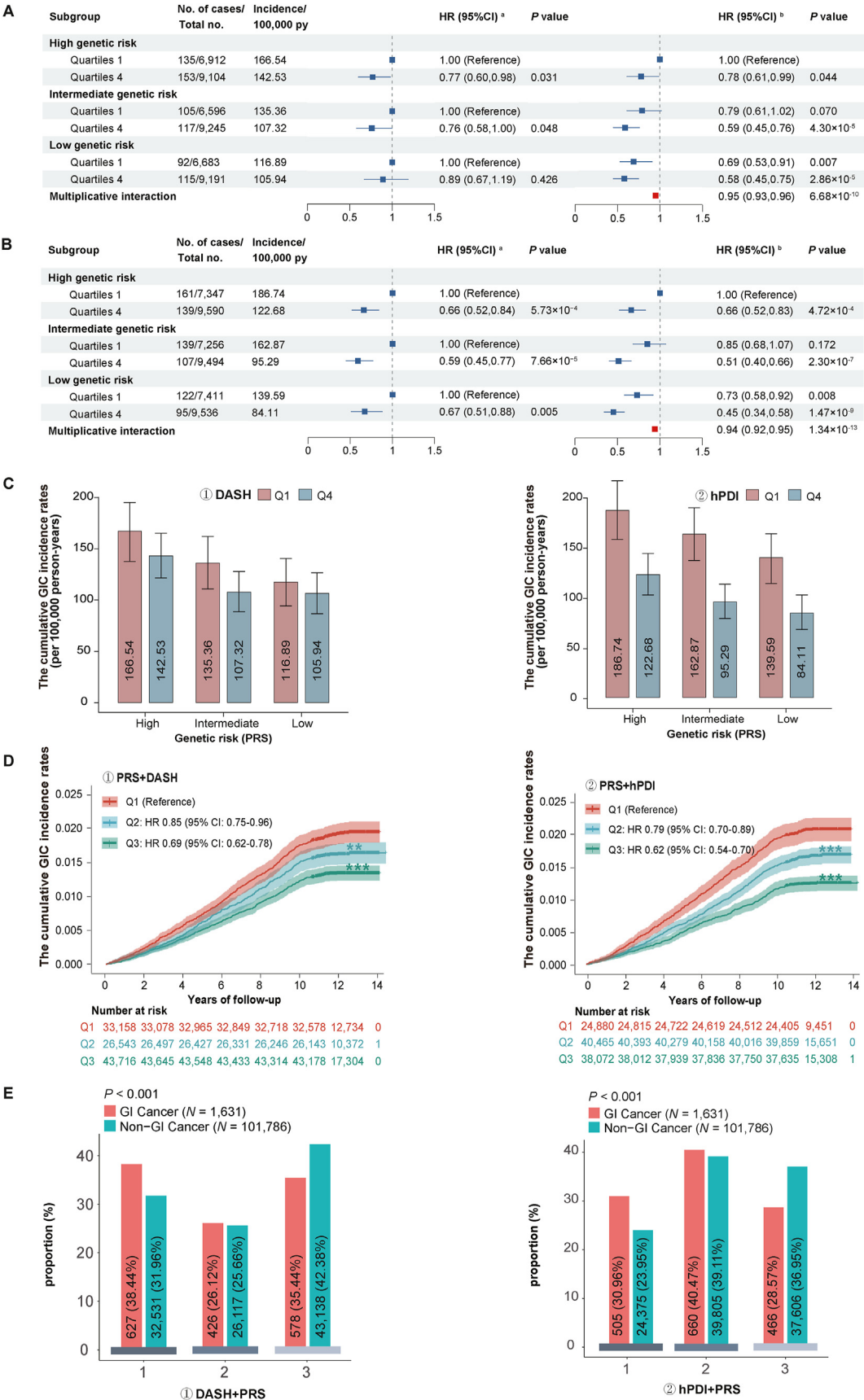
college or university degree, current smoking status, heavy alcohol consumption, and no family history. The association between hPDI and GIC risk was observed in all subgroups by age, gender, education level, smoking status, and obesity. However, when stratified by alcohol consumption and family history of GIC, there remained significant correlations only in moderate drinkers, heavy drinkers, and participants without family history (Supplementary Figure 2B). In the lag analyses (Supplementary Table 7), the association between hPDI and GIC risk remained when we included a lag of increasing duration (2, 4, 6, and 8 y). Participants with the highest quartile of DASH score had a lower risk of GIC in a 2-y lagged analysis, but no statistically significant association was noted with longer lag periods. Sensitive analysis showed that after excluding participants with BMI below the 1st or above the 99th centiles, association effects of DASH, hPDI, and GIC were generally consistent with those from main analysis (Supplementary Table 8).

Association of PRS with GIC risk

Considering GIC incidence may vary among people at different genetic risk, we constructed a GIC-PRS and observed that GIC risk associated with genetic susceptibility decreased in a dose-response manner (P -trend = 6.49×10^{-9}) after adjusting all confounders in model 3 (Table 2). The risk of GIC decreased to 0.80-fold and 0.70-fold in those at intermediate (95% CI: 0.72, 0.90; $P = 2.15 \times 10^{-4}$) and low (95% CI: 0.62, 0.79; $P = 9.43 \times 10^{-9}$) genetic risk, respectively, compared with participants at high-genetic risk. Also, similar results were found in model 1 without adjustment (P -trend = 2.24×10^{-8}) and model 2 adjusted for age and gender (P -trend = 7.12×10^{-9}) (Table 2). Furthermore, restricted cubic spline regression also characterized a clear dose-response association between PRS and GIC risk, which tended to be linear rather than nonlinear (P -nonlinearity = 0.612; Figure 1D). In addition, the cumulative risk of developing GIC was gradually decreased in association with lower PRS, suggesting that the GIC-PRS could achieve efficient risk stratification for GIC population (Figure 1E). When assessing the proportion of GIC population and non-GIC population within each category of genetic risk, we also found that the GIC-PRS could yield favorable discrimination. Participants with incident GIC tended to have a higher PRS than those without incident GIC (Figure 1F).

Joint effect and interaction of dietary patterns and PRS on GIC risk

We next explored the effects of healthy dietary patterns on GIC risk in each genetic risk category and analyzed the joint effect and interaction of dietary patterns and PRS on GIC risk (Figure 2A, B). First, using a low-DASH score as a reference, we observed that a high-DASH score reduced GIC risk in intermediate- and high-PRS groups, with HR of 0.76 (95 % CI: 0.58, 1.00) and 0.77 (95 % CI: 0.60, 0.98), respectively (Figure 2A). A joint analysis of DASH and genetic risk further revealed that GIC risk decreased with both elevation of dietary score and reduction of genetic predisposition. Compared with participants at high-genetic risk and low-DASH score, those at low-genetic risk and high-DASH score showed the lowest risk of GIC (HR: 0.58; 95% CI: 0.45, 0.75; $P = 2.86 \times 10^{-5}$). Significant multiplicative interaction was also found between DASH and genetic risk (HR: 0.95; 95% CI: 0.93, 0.96; P -interaction = 6.68×10^{-10}) (Figure 2A). Similar results can be seen in the analysis for hPDI (Figure 2B). Stratified analyses according to genetic risk showed high hPDI decreased GIC risk significantly in all genetic risk categories. The joint effect of low-genetic risk and high hPDI significantly reduced



(caption on next page)

GIC risk, with HR of 0.45 (95 % CI: 0.34, 0.58). We also observed a significant multiplicative interaction effect between hPDI and genetic risk (HR: 0.94; 95% CI: 0.92, 0.95; P -interaction = 1.34×10^{-11}) (Figure 2B).

Furthermore, within each category of genetic risk, the degrees of dietary score was associated with a decrease in the cumulative GIC incidence rate (Figure 2C). It is also noticed that participants with high-dietary score at low-genetic risk had a lower rate (105.94 per 100,000 person-years for DASH, 84.11 per 100,000 person-years for hPDI) of GIC compared with that (166.54 per 100,000 person-years for DASH, 186.74 per 100,000 person-years for hPDI) of those with low-dietary score at high-genetic risk (Figure 2C). To further validate the combined effect of dietary patterns and PRS on GIC risk, we assessed the cumulative GIC incidence rates, and the risk gradients during the follow-up also confirmed the lower GIC risk in participants with high-dietary scores and low-genetic risk (Figure 2D). Additionally, the integration values of dietary patterns and PRS enabled a favorable distinction between GIC population and non-GIC population. Non-GIC population consisted of a higher proportion of people with healthy dietary scores together with low PRS (Figure 2E). These findings suggest that genetic predisposition is not deterministic for GIC incidence, and maintaining a healthy dietary pattern may provide an effective way to counterbalance the genetic risk of GIC.

Discussion

This large-scale cohort study aimed to investigate how dietary patterns and genetic susceptibility contribute to the risk of GIC independently and jointly among 105,463 UK Biobank participants. Our findings indicated that greater adherence to DASH or hPDI was associated with lower GIC risk over time. We also generated a composite PRS to evaluate the genetic risk of GIC and found that participants at low-genetic risk compared with those at high-genetic risk were associated with ~30% lower risk of GIC. Besides, interaction analysis suggested that individuals with low-genetic risk and high-DASH score or hPDI had a lower GIC risk compared with individuals with high-genetic risk and low-DASH score or hPDI. The results also indicated that GIC risk attributed to an individual's genetic predisposition could be offset by a healthy dietary pattern, such as hPDI. Taken together, these findings greatly contribute to evidence-based prevention strategies for GIC.

Assessing diet through dietary patterns rather than a single food or nutrient has become prevalent in cancer researches because potential synergistic effects among foodstuffs and nutrients can be truly taken into account [28]. DASH and hPDI are both popular dietary patterns characterized by high intake of plant foods (e.g., fruits, vegetables, grains, nuts, and legumes) and low intake of red or processed meats and sweetened beverages. Several studies have reported inverse associations between DASH or hPDI and risk of site-specific GICs [29–34]. A prospective study of 190,949 participants from the Multiethnic Cohort

Study showed that adherence to the DASH diet was associated with reduced risk of CRC [35]. Another cohort study from the Nurses' Health Study, Nurses' Health Study II, and Health Professionals Follow-up Study found that hPDI was associated with a lower risk of total digestive system cancer as well as individual cancers in the gastrointestinal tract and the accessory organs [36]. Recently, a comparative study also indicated that adherence to 8 healthy diet patterns was associated with a 7%–13% lower risk of digestive system cancer, whereas the inverse association with gastrointestinal tract cancer was shown for all pattern scores except Alternate Mediterranean Diet and the hPDI [37]. Compared with previous studies, we confirmed positive effect of DASH and hPDI on the risk of GIC in our findings, supporting the public efforts in promoting healthy dietary patterns for everyone to reduce GIC burden [38]. In addition, our results of subgroup analyses underline that for GIC prevention, DASH is beneficial to males, the older, the high educated, current smokers, heavy drinkers, and people without family history of GIC, whereas the positive effect of hPDI is evident in almost all participants. This finding provides more new insights into the distinct benefits of dietary patterns among different subgroups of the general population, which also need further validation in future studies conducted on specific populations.

There are several possible biological mechanisms through which adherence to DASH or hPDI affect risk of GIC. Prior investigations have shown that vegetables and fruits are good sources of anticarcinogenic agents, such as β -carotene, vitamins, α -tocopherol, retinoids, phytoestrogens, and folate, which may prevent cancer by prohibiting cell growth, normalizing DNA synthesis and methylation, inducing detoxification enzymes, and reducing oxidative stress and inflammation [39,40]. High levels of dietary fiber and polyphenols of plant foods are known to reduce the risk of GIC by the production of short-chain fatty acids, reduction of fecal transit time, and improvements in insulin resistance [41,42]. Besides, low consumption of red and processed meats might further help explaining the protective association in our study. The effect of red and processed meat in incidence of GIC is related to cooking, processing, or preservation that could produce carcinogens and mutagens, including high amounts of N-nitroso compounds, heterocyclic amines, and polycyclic aromatic hydrocarbons [43]. The above potential mechanisms underlying the association between healthy dietary patterns and GIC risk further support the epidemiologic evidence found in our population-based cohort study and reveal that our results have strong biological plausibility.

With the development of genomics and biobanks, PRS has been used to evaluate genetic risk of diseases and identify individuals at an increased risk for prevention [44]. Using PRS to assess genetic risk, a growing number of studies investigated the combined effect of dietary patterns and genetic risk on poor health outcome, including obesity [45], cardiovascular disease [46], and type 2 diabetes [47]. Our findings of linear dose-response association between the 129-SNPs PRS and risk of GIC are aligned with the trend for overall or site-specific PRS in the previous studies of GIC [48–50]. Inherited susceptibility plays an

FIGURE 2. Joint effect and interaction of dietary patterns and genetic risk on GIC incidence. (A) Joint effect and interaction of DASH and genetic risk on GIC incidence. (B) Joint effect and interaction of hPDI and genetic risk on GIC incidence. (C) Risk of incident GIC according to dietary patterns and genetic factors. Error bars are 95% CIs. (D) Cumulative GIC incidences for different dietary scores within PRS. (E) Proportions of GI cancer and non-GI cancer were discriminated by different dietary scores within PRS. All HRs and 95% CIs were estimated using Cox proportional hazard models with adjustment for age, gender, ethnicity, Townsend deprivation index, education level, smoking status, alcohol consumption, BMI, physical activity, energy intake, family history of GIC, history of hypertension, heart diseases, and diabetes. ^a Participants with the lowest quarter of dietary score in each genetic risk stratification were used as the reference, respectively. ^b Participants with the lowest quarter of dietary score and high-genetic risk as the reference. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. CI, confidence interval; DASH, dietary approaches to stop hypertension; GIC, gastrointestinal cancer; hPDI, healthful plant-based diet index; HR, hazard ratio; PRS, polygenic risk score.

important role in the development of GIC and should be paid more attention to. The PRS we generated could be utilized as a measurable indicator to help raise people's awareness of their inherited susceptibility to GIC and facilitate clinical or personal decision making of risk management.

With the help of GIC-PRS, we also added to knowledge on the association of the combination of dietary patterns and genetic susceptibility with GIC risk, which indicated that the excess risks of gene-related GIC may partly result from the dietary exposure of participants. For high-genetic risk population, the incidence risk was significantly associated with 23% and 34% reductions by adhering DASH and hPDI, respectively. Thus, it is necessary to keep healthy dietary patterns to prevent GIC, because a healthy dietary pattern could partly counteract the negative influence of genetic background. The benefits also extend to people at low-genetic risk, further suggesting that maintaining healthy dietary patterns would be a key intervention for improving gastrointestinal health in general population. Apart from that, the joint analyses revealed that adherence to DASH or hPDI and low-genetic risk acted synergistically in relation to the decreased risk of GIC. The combined effect of low-genetic predisposition and high-dietary score was stronger than the product of their individual effects. In the clinical practice of precision medicine, integrating dietary patterns and genetic risk could provide more precise evaluation of individual GIC risk. However, the precise mechanisms underlying the observed interactions remain uncertain. Epigenetic mechanisms may partly explain the modifying effect of healthy dietary patterns on genetic predisposition to GIC. Several diets high in polyphenols have been associated with the attenuation of tumor progression and prevention of metastasis through influencing DNA methylation, histone modifications, and noncoding RNA expression in cancer cells [8]. The enormous potential of healthy dietary patterns to reverse adverse epigenetic marks in cancer cells requires more functional studies in the future.

Based on the evidence from this large-scale cohort study, dietary interventions focusing on healthy dietary patterns can be considered as a targeted approach for GIC prevention, especially for people at higher GIC genetic risk who would benefit more. In the future, GIC prevention work should attach great importance to improving modifiable risk factors, such as diet, strengthening related health education, and instructing individuals to comply with healthy dietary patterns, such as DASH and hPDI. Improving diet quality under genetic risk stratification provides a more achievable way to alleviate GIC burden.

Strengths and limitations

The strengths of our study are the large-scale sample with a prospective design, long-term follow-up, extensive assessments of dietary intake, as well as detailed information on multiple confounders. We adjusted for potential confounders and performed sensitivity analyses, demonstrating the robustness of our findings. Moreover, we analyzed dietary scores instead of individual nutrients or foods in our study, which took into account the complex and multidimensional nature of the diet. Additionally, the inclusion of genetic data in the UK Biobank allowed us to construct the specific PRS to predict GIC risk and take it into analysis. However, our study also has several limitations. First, dietary information was collected through 24-h dietary recall, which might be subject to recall error and misreporting. Second, although we have adjusted for a wide range of confounders, there may still exist residual or unmeasured confounding, such as medication use. Last but not least, it is necessary to conduct further studies to examine the impact of GIC heterogeneity including subtypes, localizations,

microsatellite unstable (MSI) status, and tumor node metastasis (TNM) stage on the relationship between dietary patterns and GIC incidence. MSI defines the typical genomic phenotype of DNA repeat region insertion and deletion mutations [51], which is considered an important factor in the occurrence and development of tumors. TNM is the most commonly used staging system for gastrointestinal cancer. Determining MSI status and TNM stage plays a critical role in achieving a more comprehensive and profound understanding of gene-diet interaction on GIC.

Conclusions

This prospective cohort study of 105,463 UK adults provided evidence that higher DASH or hPDI scores and lower PRS were associated with lower GIC risk independently and jointly. Adherence to DASH or hPDI may partly offset the genetic predisposition to GIC. These findings lend robust support to the development of precision medicine strategies that consider both dietary patterns and genetics to ameliorate gastrointestinal health.

Acknowledgments

We thank all participants and researchers from the UK Biobank. This research has been conducted using the UK Biobank Resource under Application Number 94939.

Author contributions

The authors' responsibilities were as follows – YC, CH, JH, LF, XX, JX: conceived and designed the study; YC, CH, XX: performed data acquisition; YC, CH, JH, LF, XX, JX: performed statistical analysis and interpretation of data; YC, CH, JH, LF, XX, JX: drafted the manuscript; YW, YZ, JT, XZ, MJ, XM: critically revised the manuscript; JT, XZ, MJ, XM: provided administrative, technical, or material support; JT, XZ, MJ, XM: had primary responsibility for final content. All authors read and approved the final manuscript.

Conflict of interest

The authors report no conflicts of interest.

Funding

This work was supported by the Key Program of National Natural Science Foundation of China (NSFC-82130098) for Xiaoping Miao; the Fundamental Research Funds for the Central Universities (WHU:2042022kf1205) and the Knowledge Innovation Program of Wuhan (whxksj011) for Jianbo Tian.

Data availability

Data used in this study are available on application to the UK Biobank (www.ukbiobank.ac.uk/register-apply). Analytical methods and study materials will be available to other researchers from the corresponding authors upon reasonable request for the purposes of reproducing the results or replicating the procedure.

Ethics approval

The North West Multi-Centre Research Ethics Committee approved the collection and use of UK Biobank data. All participants provided written informed consent. Institutional review board approval was waived for this analysis because of the publicly available and deidentified data. The study was conducted under the approval of the Institutional Review Board of Wuhan University (ID: 2022042).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ajcnut.2023.11.015>.

References

- [1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA Cancer J. Clin.* 71 (3) (2021) 209–249, <https://doi.org/10.3322/caac.21660>.
- [2] W.Q. Li, Y. Park, J.W. Wu, J.S. Ren, A.M. Goldstein, P.R. Taylor, et al., Index-based dietary patterns and risk of esophageal and gastric cancer in a large cohort study, *Clin. Gastroenterol. Hepatol.* 11 (9) (2013) 1130–1136.e2, <https://doi.org/10.1016/j.cgh.2013.03.023>.
- [3] A. Agudo, V. Cayssials, C. Bonet, A. Tjonneland, K. Overvad, M.C. Boutron-Ruault, et al., Inflammatory potential of the diet and risk of gastric cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) study, *Am. J. Clin. Nutr.* 107 (4) (2018) 607–616, <https://doi.org/10.1093/ajcn/nqy002>.
- [4] H. Steel, S.Y. Park, T. Lim, D.O. Stram, C.J. Boushey, J.R. Hebert, et al., Diet quality and pancreatic cancer incidence in the multiethnic cohort, *Cancer Epidemiol. Biomarkers Prev.* 32 (1) (2023) 123–131, <https://doi.org/10.1158/1055-9965.EPI-22-0564>.
- [5] B. Wu, R.L. Zhou, Q.J. Ou, Y.M. Chen, Y.J. Fang, C.X. Zhang, Association of plant-based dietary patterns with the risk of colorectal cancer: a large-scale case-control study, *Food Funct.* 13 (20) (2022) 10790–10801, <https://doi.org/10.1039/d2fo01745h>.
- [6] P. Wang, M. Song, A.H. Eliassen, M. Wang, E.L. Giovannucci, Dietary patterns and risk of colorectal cancer: a comparative analysis, *Int. J. Epidemiol.* 52 (1) (2023) 96–106, <https://doi.org/10.1093/ije/dyac230>.
- [7] S. Moazzen, K.W.J. van der Sloot, R.J. Vonk, G.H. de Bock, B.Z. Alizadeh, Diet quality and upper gastrointestinal cancers risk: a meta-analysis and critical assessment of evidence quality, *Nutrients* 12 (6) (2020) 1863, <https://doi.org/10.3390/nu12061863>.
- [8] S.E. Steck, E.A. Murphy, Dietary patterns and cancer risk, *Nat. Rev. Cancer.* 20 (2) (2020) 125–138, <https://doi.org/10.1038/s41568-019-0227-4>.
- [9] A.J. MacFarlane, P.J. Stover, Convergence of genetic, nutritional and inflammatory factors in gastrointestinal cancers, *Nutr. Rev.* 65 (12 Pt 2) (2007) S157–S166, <https://doi.org/10.1111/j.1753-4887.2007.tb00355.x>.
- [10] J.R. Huyghe, S.A. Bien, T.A. Harrison, H.M. Kang, S. Chen, S.L. Schmit, et al., Discovery of common and rare genetic risk variants for colorectal cancer, *Nat. Genet.* 51 (1) (2019) 76–87, <https://doi.org/10.1038/s41588-018-0286-6>.
- [11] A.T. Kunzmann, M. Cañadas Garre, A.P. Thrift, U.C. McMenamin, B.T. Johnston, C.R. Cardwell, et al., Information on genetic variants does not increase identification of individuals at risk of esophageal adenocarcinoma compared to clinical risk factors, *Gastroenterology* 156 (1) (2019) 43–45, <https://doi.org/10.1053/j.gastro.2018.09.038>.
- [12] W. Liu, T. Wang, M. Zhu, G. Jin, Healthy diet, polygenic risk score, and upper gastrointestinal cancer risk: a prospective study from UK Biobank, *Nutrients* 15 (6) (2023) 1344, <https://doi.org/10.3390/nu15061344>.
- [13] C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, et al., UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age, *PLoS Med* 12 (3) (2015) e1001779, <https://doi.org/10.1371/journal.pmed.1001779>.
- [14] C. Bycroft, C. Freeman, D. Petkova, G. Band, L.T. Elliott, K. Sharp, et al., The UK Biobank resource with deep phenotyping and genomic data, *Nature* 562 (7726) (2018) 203–209, <https://doi.org/10.1038/s41586-018-0579-z>.
- [15] B. Liu, H. Young, F.L. Crowe, V.S. Benson, E.A. Spencer, T.J. Key, et al., Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies, *Public Health Nutr* 14 (11) (2011) 1998–2005, <https://doi.org/10.1017/S1368980011000942>.
- [16] A. Perez-Cornago, Z. Pollard, H. Young, M. van Uden, C. Andrews, C. Piernas, et al., Description of the updated nutrition calculation of the Oxford WebQ questionnaire and comparison with the previous version among 207,144 participants in UK Biobank, *Eur. J. Nutr.* 60 (7) (2021) 4019–4030, <https://doi.org/10.1007/s00394-021-02558-4>.
- [17] C. Piernas, A. Perez-Cornago, M. Gao, H. Young, Z. Pollard, A. Mulligan, et al., Describing a new food group classification system for UK biobank: analysis of food groups and sources of macro- and micronutrients in 208,200 participants, *Eur. J. Nutr.* 60 (5) (2021) 2879–2890, <https://doi.org/10.1007/s00394-021-02535-x>.
- [18] K.E. Bradbury, H.J. Young, W. Guo, T.J. Key, Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire, *J. Nutr. Sci.* 7 (2018) e6, <https://doi.org/10.1017/jns.2017.66>.
- [19] A.R. Curtis, K.M. Livingstone, R.M. Daly, B. Brayner, G. Abbott, N. Kiss, Dietary patterns, malnutrition, muscle loss and sarcopenia in cancer survivors: findings from the UK Biobank, *J. Cancer Surviv.* (2023), <https://doi.org/10.1007/s11764-023-01428-8>.
- [20] T.T. Fung, S.E. Chiuve, M.L. McCullough, K.M. Rexrode, G. Logroscino, F.B. Hu, Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women, *Arch. Intern. Med.* 168 (7) (2008) 713–720, <https://doi.org/10.1001/archinte.168.7.713>.
- [21] A. Satija, S.N. Bhupathiraju, E.B. Rimm, D. Spiegelman, S.E. Chiuve, L. Borgi, et al., Plant-based dietary patterns and incidence of type 2 diabetes in US men and women: results from three prospective cohort studies, *PLoS Med* 13 (6) (2016) e1002039, <https://doi.org/10.1371/journal.pmed.1002039>.
- [22] C.A. Gonzalez, G. Pera, A. Agudo, H.B. Bueno-de-Mesquita, M. Ceroti, H. Boeing, et al., Fruit and vegetable intake and the risk of stomach and oesophagus adenocarcinoma in the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST), *Int. J. Cancer.* 118 (10) (2006) 2559–2566, <https://doi.org/10.1002/ijc.21678>.
- [23] A.P. Klein, B.M. Wolpin, H.A. Risch, R.Z. Stolzenberg-Solomon, E. Mocci, M. Zhang, et al., Genome-wide meta-analysis identifies five new susceptibility loci for pancreatic cancer, *Nat. Commun.* 9 (1) (2018) 556, <https://doi.org/10.1038/s41467-018-02942-5>.
- [24] J. Euesden, C.M. Lewis, P.F. O'Reilly, PRSice: polygenic risk score software, *Bioinformatics* 31 (9) (2015) 1466–1468, <https://doi.org/10.1093/bioinformatics/btu848>.
- [25] A.P. Keszei, L.J. Schouten, R.A. Goldbohm, P.A. van den Brandt, Red and processed meat consumption and the risk of esophageal and gastric cancer subtypes in the Netherlands Cohort Study, *Ann. Oncol.* 23 (9) (2012) 2319–2326, <https://doi.org/10.1093/annonc/mdr615>.
- [26] C.L. Craig, A.L. Marshall, M. Sjostrom, A.E. Bauman, M.L. Booth, B.E. Ainsworth, et al., International physical activity questionnaire: 12-country reliability and validity, *Med. Sci. Sports Exerc.* 35 (8) (2003) 1381–1395, <https://doi.org/10.1249/01.MSS.0000078924.61453.FB>.
- [27] J. Adams, V. Ryan, M. White, How accurate are Townsend Deprivation Scores as predictors of self-reported health? A comparison with individual level data, *J. Public Health (Oxf.)* 27 (1) (2005) 101–106, <https://doi.org/10.1093/pubmed/fdh193>.
- [28] E.M. Cespedes, F.B. Hu, Dietary patterns: from nutritional epidemiologic analysis to national guidelines, *Am. J. Clin. Nutr.* 101 (5) (2015) 899–900, <https://doi.org/10.3945/ajcn.115.110213>.
- [29] F. Toorang, B. Sasanfar, M. Hadji, A. Esmaillzadeh, K. Zendeheidi, Adherence to "dietary approaches to stop hypertension" eating plan in relation to gastric cancer, *Nutr. J.* 19 (1) (2020) 40, <https://doi.org/10.1186/s12937-020-00560-w>.
- [30] S. Julian-Serrano, J. Reedy, K. Robien, R. Stolzenberg-Solomon, Adherence to 5 diet quality indices and pancreatic cancer risk in a large US prospective cohort, *Am. J. Epidemiol.* 191 (9) (2022) 1584–1600, <https://doi.org/10.1093/aje/kwac082>.
- [31] H. Tangestani, A. Salari-Moghaddam, H. Ghalandari, H. Emamat, Adherence to the dietary approaches to stop hypertension (DASH) dietary pattern reduces the risk of colorectal cancer: a systematic review and meta-analysis, *Clin. Nutr.* 39 (10) (2020) 2975–2981, <https://doi.org/10.1016/j.clnu.2020.02.002>.
- [32] N.W. Solomons, Plant-based diets are traditional in developing countries: 21st century challenges for better nutrition and health, *Asia Pac. J. Clin. Nutr.* 9 (Suppl 1) (2000) S41–S54, <https://doi.org/10.1046/j.1440-6047.2000.00165.x>.
- [33] Y. Zhao, J. Zhan, Y. Wang, D. Wang, The relationship between plant-based diet and risk of digestive system cancers: a meta-analysis based on 3,059,009 subjects, *Front. Public Health* 10 (2022) 892153, <https://doi.org/10.3389/fpubh.2022.892153>.
- [34] J. Kim, C.J. Boushey, L.R. Wilkens, C.A. Haiman, L. Le Marchand, S.Y. Park, Plant-based dietary patterns defined by a priori indices and colorectal cancer risk by sex and race/ethnicity: the multiethnic cohort study, *BMC Med* 20 (1) (2022) 430, <https://doi.org/10.1186/s12916-022-02623-7>.
- [35] S. Park, C.J. Boushey, L.R. Wilkens, C.A. Haiman, L. Le Marchand, High-quality diets associate with reduced risk of colorectal cancer: analyses of diet quality indexes in the multiethnic cohort, *Gastroenterology* 153 (2) (2017) 386–394.e2, <https://doi.org/10.1053/j.gastro.2017.04.004>.
- [36] J. Kim, J. Khil, H. Kim, N. Keum, X. Zhang, E. Giovannucci, Plant-based dietary patterns and the risk of digestive system cancers in 3 large prospective

- cohort studies, *Eur. J. Epidemiol.* 38 (6) (2023) 617–627, <https://doi.org/10.1007/s10654-023-01007-2>.
- [37] J. Kim, Y. Zhang, H. Kim, Y. Zhang, X. Zhang, E. Giovannucci, A comparative study of healthy dietary patterns for incident and fatal digestive system cancer, *Am. J. Gastroenterol.* 118 (11) (2023) 2061–2070, <https://doi.org/10.14309/ajg.0000000000002448>.
- [38] J.H. Bi, H.Y. Yuan, Y. Jiang, Y. Zhang, W.W. Zheng, L. Zhang, et al., Incidence, mortality features and lifetime risk estimation of digestive tract cancers in an urban district of Shanghai, China, *J. Epidemiol. Glob. Health* 12 (3) (2022) 248–257, <https://doi.org/10.1007/s44197-022-00047-3>.
- [39] K.A. Steinmetz, J.D. Potter, Vegetables, fruit, and cancer. II. Mechanisms, *Cancer Causes Control* 2 (6) (1991) 427–442, <https://doi.org/10.1007/BF00054304>.
- [40] J.M. Lu, P.H. Lin, Q. Yao, C. Chen, Chemical and molecular mechanisms of antioxidants: experimental approaches and model systems, *J. Cell. Mol. Med.* 14 (4) (2010) 840–860, <https://doi.org/10.1111/j.1582-4934.2009.00897.x>.
- [41] D. Aune, D.S. Chan, R. Lau, R. Vieira, D.C. Greenwood, E. Kampman, et al., Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies, *BMJ* 343 (2011) d6617, <https://doi.org/10.1136/bmj.d6617>.
- [42] M. Song, A.T. Chan, J. Sun, Influence of the gut microbiome, diet, and environment on risk of colorectal cancer, *Gastroenterology* 158 (2) (2020) 322–340, <https://doi.org/10.1053/j.gastro.2019.06.048>.
- [43] A.N. Samraj, O.M. Pearce, H. Laubli, A.N. Crittenden, A.K. Bergfeld, K. Banda, et al., A red meat-derived glycan promotes inflammation and cancer progression, *Proc. Natl. Acad. Sci. U S A* 112 (2) (2015) 542–547, <https://doi.org/10.1073/pnas.1417508112>.
- [44] A. Torkamani, N.E. Wineinger, E.J. Topol, The personal and clinical utility of polygenic risk scores, *Nat. Rev. Genet.* 19 (9) (2018) 581–590, <https://doi.org/10.1038/s41576-018-0018-x>.
- [45] Y. Heianza, T. Zhou, D. Sun, F.B. Hu, L. Qi, Healthful plant-based dietary patterns, genetic risk of obesity, and cardiovascular risk in the UK Biobank study, *Clin. Nutr.* 40 (7) (2021) 4694–4701, <https://doi.org/10.1016/j.clnu.2021.06.018>.
- [46] K.M. Livingstone, G. Abbott, S.J. Bowe, J. Ward, C. Milte, S.A. McNaughton, Diet quality indices, genetic risk and risk of cardiovascular disease and mortality: a longitudinal analysis of 77004 UK Biobank participants, *BMJ Open* 11 (4) (2021) e045362, <https://doi.org/10.1136/bmjopen-2020-045362>.
- [47] J. Merino, M. Guasch-Ferre, J. Li, W. Chung, Y. Hu, B. Ma, et al., Polygenic scores, diet quality, and type 2 diabetes risk: an observational study among 35,759 adults from 3 US cohorts, *PLoS Med* 19 (4) (2022) e1003972, <https://doi.org/10.1371/journal.pmed.1003972>.
- [48] Y. Liu, C. Yan, S. Yin, T. Wang, M. Zhu, L. Liu, et al., Genetic risk, metabolic syndrome, and gastrointestinal cancer risk: a prospective cohort study, *Cancer Med* 12 (1) (2023) 597–605, <https://doi.org/10.1002/cam4.4923>.
- [49] G. Jin, J. Lv, M. Yang, M. Wang, M. Zhu, T. Wang, et al., Genetic risk, incident gastric cancer, and healthy lifestyle: a meta-analysis of genome-wide association studies and prospective cohort study, *Lancet. Oncol.* 21 (10) (2020) 1378–1386, [https://doi.org/10.1016/S1470-2045\(20\)30460-5](https://doi.org/10.1016/S1470-2045(20)30460-5).
- [50] A.N. Archambault, Y.R. Su, J. Jeon, M. Thomas, Y. Lin, D.V. Conti, et al., Cumulative burden of colorectal cancer-associated genetic variants is more strongly associated with early-onset vs late-onset cancer, *Gastroenterology* 158 (5) (2020) 1274–1286.e12, <https://doi.org/10.1053/j.gastro.2019.12.012>.
- [51] S.N. Thibodeau, G. Bren, D. Schaid, Microsatellite instability in cancer of the proximal colon, *Science* 260 (5109) (1993) 816–819, <https://doi.org/10.1126/science.8484122>.