

Healthy lifestyles, genetic modifiers, and colorectal cancer risk: a prospective cohort study in the UK Biobank

Jungyoon Choi, Guochong Jia, Wanqing Wen, Xiao-Ou Shu, and Wei Zheng

Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN

ABSTRACT

Background: Both genetic and lifestyle factors play an etiologic role in colorectal cancer (CRC).

Objectives: We evaluated potential gene–environment interactions in CRC risk.

Methods: We used data from 346,297 participants in the UK Biobank cohort. Healthy lifestyle scores (HLSs) were constructed using 8 lifestyle factors, primarily according to the American Cancer Society guidelines, and were categorized into unhealthy, intermediate, and healthy groups. A polygenic risk score (PRS) was created using 95 genetic risk variants identified by genome-wide association studies of CRC and was categorized by tertile. Cox models were used to estimate the HRs and 95% CIs of CRC risk associated with the HLS and PRS.

Results: During a median follow-up of 5.8 y, 2066 incident cases of CRC were identified. Healthier HLSs were associated with reduced risk of CRC in a dose–response manner. The risk reduction was more apparent among those with high PRS (HR_{healthy vs. unhealthy HLS1}: 0.58; 95% CI: 0.43, 0.79 for men and 0.71; 0.58, 0.85 for men and women combined) than those with low PRS. Although no multiplicative interactions were identified, the HLS1 and PRS showed a significant additive interaction ($P = 0.02$ for all participants combined, 0.04 for men). In analyses including all participants, the adjusted CRC cumulative risk from age 40 to 75 y was 6.40% for those with high PRS/unhealthy HLS1, with a relative excess risk due to interaction of 0.58 (95% CI: 0.06, 1.10), compared with 2.09% among those with low PRS/healthy HLS1. This pattern was more apparent among those who reported not having received any bowel screening before baseline.

Conclusions: Although the observational nature of the study precludes proof of causality, our findings suggest that individuals with a high genetic susceptibility could benefit more substantially than those with a low genetic risk from lifestyle modification in reducing CRC risk. *Am J Clin Nutr* 2021;113:810–820.

Keywords: colorectal cancer, gene–environment interactions, polygenic risk scores, healthy lifestyle factors, epidemiology

Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed malignancy in the world (1). Both genetic and lifestyle

factors are known to play an etiologic role in this common cancer. Individuals with pathogenic mutations in CRC susceptibility genes are at a very high risk of developing this cancer. However, germline mutations of these genes are rare, accounting for <10% of CRC cases in the general population. Recent genome-wide association studies (GWAS) have identified a large number of common genetic variants associated with CRC risk. Although each of these risk variants confers a small to moderate risk of CRC, the polygenic risk score (PRS), a measure of the cumulative effect of these variants, is strongly associated with CRC risk (2–5).

Multiple lifestyle factors have been reported to be associated with CRC risk. These risk factors include obesity (6, 7), high intake of red and processed meat (8, 9), excessive alcohol consumption (10), and tobacco smoking (11). Reduction in these risk factors, alone or in combination, has been shown to be associated with a significantly reduced risk of CRC (12, 13). Recent studies have suggested that healthy lifestyles may reduce the risk of CRC, regardless of an individual’s genetic risk (14, 15). However, it is unclear whether there is an interaction between genetic factors and modifiable lifestyle factors, in particular whether the genetic risk can be mitigated by healthy lifestyles. We sought to answer these questions using data from a large prospective cohort study conducted in the UK Biobank.

Supported in part by funds provided by the Anne Potter Wilson chair endowment at Vanderbilt University and a research grant (R01 CA188214) from the US National Cancer Institute. This research was conducted using the UK Biobank Resource (application number 40685).

Supplemental Tables 1–6 and Supplemental Figures 1–5 are available from the “Supplementary data” link in the online posting of the article and from the same link in the online table of contents at <https://academic.oup.com/ajcn/>.

Address correspondence to WZ (e-mail: wei.zheng@vanderbilt.edu).

Abbreviations used: ACS, American Cancer Society; CRC, colorectal cancer; GWAS, genome-wide association studies; HLS, healthy lifestyle score; ICD, International Classification of Diseases; LD, linkage disequilibrium; NHS, National Health Service; PRS, polygenic risk score; RERI, relative excess risk due to interaction; SNP, single-nucleotide polymorphism; UKBB, UK Biobank Axiom; UKBL, UK BiLEVE Axiom; WHR, waist-to-hip ratio. Received July 13, 2020. Accepted for publication December 1, 2020.

First published online March 1, 2021; doi: <https://doi.org/10.1093/ajcn/nqaa404>.

Methods

Study population

The UK Biobank is a population-based cohort study which has recruited >500,000 adults across England, Scotland, and Wales. The design and methods of the UK Biobank study have been previously described (16). At enrollment, information on the participants' sociodemographic characteristics, health and medical history, and diet and other lifestyle factors was collected using a self-administered touchscreen questionnaire (<http://www.ukbiobank.ac.uk>) and nurse-led interviews. During the interviews, the participants' body weight, waist and hip circumferences, and height were measured by trained staff using standardized procedures. Data on bowel screening including tests for blood in the stool/feces or a colonoscopy or a sigmoidoscopy were obtained at the baseline survey.

Data on the diagnosis of site-specific incident cancers were provided by the National Health Service (NHS) Information Centre for participants from England and Wales (follow-up through 31 March 2016) and by the NHS Central Register Scotland for participants from Scotland (follow-up through 31 October 2015). Cancers were coded by use of the International Classification of Diseases, 9th Revision (ICD-9), or the International Classification of Diseases, 10th Revision (ICD-10). The outcome for this study is incident CRC as the first cancer diagnosis, with codes of ICD-9 153, 154.0, and 154.1 or ICD-10 C18, C19, and C20. The study was approved by the relevant ethical committees for the UK Biobank and Vanderbilt University Medical Center. All participants provided written informed consent.

Construction of healthy lifestyle scores

We created healthy lifestyle scores (HLSs) primarily based on the American Cancer Society (ACS) Guidelines on Nutrition and Physical Activity for Cancer Prevention and the guideline to stay away from tobacco (17, 18). Eight variables were used in constructing the HLS: BMI (kg/m^2), waist-to-hip ratio (WHR), physical activity, sedentary time, processed and red meat intake, vegetable and fruit intake, alcohol consumption, and tobacco smoking. We created 2 scores: a met/not met guidelines score (HLS1) and a sex-specific weighted score (HLS2). **Supplemental Table 1** presents the categories for these variables.

The HLS1 was created for each participant by counting them as having met or not met ACS guidelines for each component. Each variable was given a score of 0 or 1, with 1 representing the healthy behavior category. BMI was calculated using the values for weight and height provided at the baseline interview and participants were classified as having met the recommendation if they were within 18.5–24.9. We included WHR in our study because strong evidence for a potential causal association of WHR with CRC risk was presented in the recent World Cancer Research Fund/American Institute for Cancer Research review (19). WHR results were classified as having met a healthy lifestyle requirement if they were <0.90 for men and <0.85 for women, according to WHO guidelines (20). Participants were classified as having met current physical activity recommendations if they reported 6–7 d/wk of moderate activity and 3–5 d/wk of vigorous activity, or 6–7 d/wk of vigorous activity, an approximation of 150–300 min/wk

of moderate-intensity or 75–150 min/wk of vigorous-intensity physical activities or an equivalent combination of moderate- and vigorous-intensity activity as recommended by the physical activity guidelines for Americans (21). For the sedentary time variable, we used time spent on watching television and using computer for leisure as no information was collected regarding other sedentary activities in the cohort. Because there were no specific recommendations regarding sedentary time, we used the lowest tertile (<3 h/d) as meeting the guideline. We generated scores for 2 dietary factors that are known to be associated with CRC risk with data collected using an FFQ (22), intakes of processed and red meat and vegetable and fruit. Consumption of processed and red meat was classified as having met the recommendation if it was <4 times per wk. Vegetable and fruit intake was classified as having met a healthy lifestyle requirement if it was >5 servings per d. A half cup (118.3 mL) was regarded as 1 serving. We classified never and seldom (special occasions only, and 1–3 times a month) alcohol drinkers as having met the recommendation. Never smokers were considered to have met the “Stay Away from Tobacco” guideline. The HLS was then constructed by taking the sum of the scores for body fat measures (BMI and WHR), physical activity, sedentary time, diet (processed and red meat, vegetable and fruit), alcohol consumption, and tobacco smoking. The HLS1 ranged from 0 to 8, and was analyzed in this study according to 3 categories: unhealthy (score 0, 1), intermediate (2, 3), and healthy (≥ 4), based on the tertile distribution of HLS1 in noncases. To capture a more detailed spectrum of each lifestyle component, we used sex-specific \log_e -HR estimates for each lifestyle factor as the weights to add all 8 lifestyle factors into a single score (HLS2) (**Supplemental Table 2**). We divided each lifestyle component into 3 categories for HR estimates. Details are presented in Supplemental Table 1.

Construction of polygenic risk score

We obtained genotype imputation data from 487,154 participants of the UK Biobank. Samples were genotyped using 2 arrays sharing a 95% marker content, the UK BiLEVE Axiom (UKBL; 807,411 markers) and the UK Biobank Axiom (UKBB; 825,927 markers). These genotyping data were imputed using reference panels of the Haplotype Reference Consortium, or UK10K, and 1000 Genomes Project phase 3. We excluded individuals marked as outliers for heterozygosity, low call rates, and sex chromosome aneuploidy ($n = 630$). European individuals were identified from the genotype data by projecting all of the UK Biobank samples on the first 2 major principal components of 4 populations included in the 1000 Genomes Project (CEU, YRI, CHB, and JPT) (23). Individuals not falling into the CEU cluster were excluded ($n = 23,409$). Self-reporting as non-European was also excluded ($n = 4916$). In the dataset from UK Biobank, a kinship coefficient was estimated for each pair of samples by using KING's robust estimator (24). We excluded second-degree (or higher) related individuals (kinship coefficient ≥ 0.0442 ; $n = 35,067$).

A genetic risk score was built based on 95 risk variants identified in previous GWAS of CRC risk (**Supplemental Table 3**) (3, 25). These variants were selected by reviewing the GWAS catalog and PubMed publications. We selected from the most recent studies with the largest sample sizes of individuals of European ancestry. Using the conventional

genome-wide significance threshold ($P < 5 \times 10^{-8}$), single-nucleotide polymorphisms (SNPs) showing an association with P values at or below this threshold were included in our study. Some risk variants with a confirmed association with the cancer of interest may not have been significant at $P < 5 \times 10^{-8}$ in the latest studies due to small sample sizes. These variants were also included in the current study to construct the PRS using regression coefficients from either the latest studies (if available) or previous studies. Cancer risk variants on the X chromosome, those reported exclusively from non-European populations, or those that were in high linkage disequilibrium (LD; $r^2 \geq 0.2$), were also excluded in this study. For some previously reported risk variants which were not available in the data from UK Biobank, SNPs in high LD ($n = 3$; $r^2 \geq 0.85$) were selected for the study. After imputation, no participant had any missing value for the 95 CRC risk variants selected for this study.

We calculated the PRS by summing the product of the weight and the number of risk alleles (0, 1, and 2) for each risk variant across all GWAS identified risk variants for all study participants. Details on the derivation of the genetic risk score have been published recently (3). The PRS was categorized as low (score: ≤ 7.6613), intermediate (7.6614–8.0567), and high (> 8.0567) based on the tertile distribution of PRS in noncases.

Analytical cohort

We included only participants of European ancestry in this study, as the PRS was derived using risk variants identified from GWAS conducted in this population. The process for construction of the analytical cohort can be found in **Figure 1**. Of the 423,374 cohort members remaining after previously mentioned exclusions, we also excluded participants who had been diagnosed with cancer at baseline ($n = 22,759$) and aged < 40 y ($n = 5$). To minimize the potential influence of preclinical diseases, we also excluded participants who had a history of inflammatory bowel diseases (Crohn's disease: ICD-9 = 555 or ICD-10 = K50, and ulcerative colitis: ICD-9 = 556 or ICD-10 = K51, $n = 1046$) (not mutually exclusive). Participants with a BMI < 18.5 kg/m² and missing data for any one component of the HLS were excluded. Missing data for education and Townsend deprivation index were set to medians. After these additional exclusions, the study population comprised 346,297 participants (161,736 men and 184,561 women). All of the missing covariates were $< 2\%$.

Statistical analyses

Comparisons of demographic characteristics of CRC cases and noncases were evaluated using chi-square tests (for categorical variables) or t -tests (for continuous variables). Our primary aim was to assess the associations between HLS and risk of CRC by PRS categories. HRs and 95% CIs for the associations of HLS and PRS with CRC risk were estimated using Cox proportional hazards models, with age as the underlying timescale left truncated at the age of baseline interview. Analyses were performed for all participants combined, and separately for men and women. Regression models were adjusted for age at enrollment, sex, education, socioeconomic status (based on the Townsend deprivation index as quintiles), bowel screening, family history of CRC, the top 5 principal components for

ancestry, and genotyping batch. The Cox models were stratified by birth cohorts, defined according to the year of birth as follows: ≤ 1941 , > 1941 to ≤ 1951 , > 1951 to ≤ 1956 , > 1956 to ≤ 1961 , > 1961 to ≤ 1966 , and > 1966 . In addition, analyses were performed separately for those who reported having or not having received a bowel screening before the baseline survey. Also, we assessed the interaction between HLS and family history of CRC in CRC risk. The assumptions of proportionality were verified using Schoenfeld residuals. To test for additive interactions between HLS and PRS, we used Cox proportional hazards regression models as previously described and estimated the relative excess risk due to interaction (RERI) and its 95% CI (26, 27). The additive interaction was assessed as to whether the estimated joint effects of the HLS and PRS were greater than the sum of the individual effect estimates for these 2 variables. To test for multiplicative interactions between the HLS and the PRS, we included an interaction term for HLS and PRS in the Cox regression models and used the Wald test to evaluate if this term was statistically significant. We also calculated the cumulative risk of CRC from age 40 to 75 y according to combinations of healthy lifestyle and genetic risk categories. We estimated CRC-free survivals by the joint distribution of PRS and HLS groups, with an adjustment of variables mentioned above. We then estimated the cumulative incidence risk of CRC as the complement of the survivals (i.e., 1 minus survivals), adjusted to the means of continuous covariates and modes of categorical variables. The UK Biobank recruited participants aged ≥ 40 y. Thus, we defined cumulative risk as the probability of developing CRC from the age of 40 to 75 y. All statistical inferences were based on 2-sided tests at a significance level of 0.05 using SAS software, version 9.4 (SAS Institute) and R v. 3.6.0 software (<https://www.r-project.org/>).

Results

During a median follow-up of 5.8 y, 2066 incident cases of CRC were identified among eligible study participants. Compared to noncases, CRC cases were more likely to be older, male, have a family history of CRC, be less educated, have had more bowel screenings, and have higher PRS (**Table 1**).

Table 2 presents the associations of each of the HLS components with risk of CRC in all participants combined and separately by sex. With a few exceptions, unhealthier behavior was, in general, associated with an increased CRC risk compared with those meeting healthy lifestyle guidelines (the reference category) for each HLS component, although not all risk estimates were statistically significant. A noticeable exception was alcohol consumption in women, for which a high intake (1–4 times per wk or almost daily) was not related to an increased risk. Given that this is a component of the ACS lifestyle guidelines, we decided to include this variable in the HLS analyses. Having a healthier HLS score was significantly related to a reduced risk of CRC in a dose–response manner (P for trend < 0.0001 for all participants combined and for men only in both analyses of HLS1 and HLS2). In women, a healthier HLS2 score was significantly related to a reduced risk of CRC (P for trend = 0.0003), while HLS1 showed a marginal significance (P for trend = 0.07). Also, there was a significant P for interaction between HLS and sex (0.007 for HLS1 and 0.01 for HLS2).

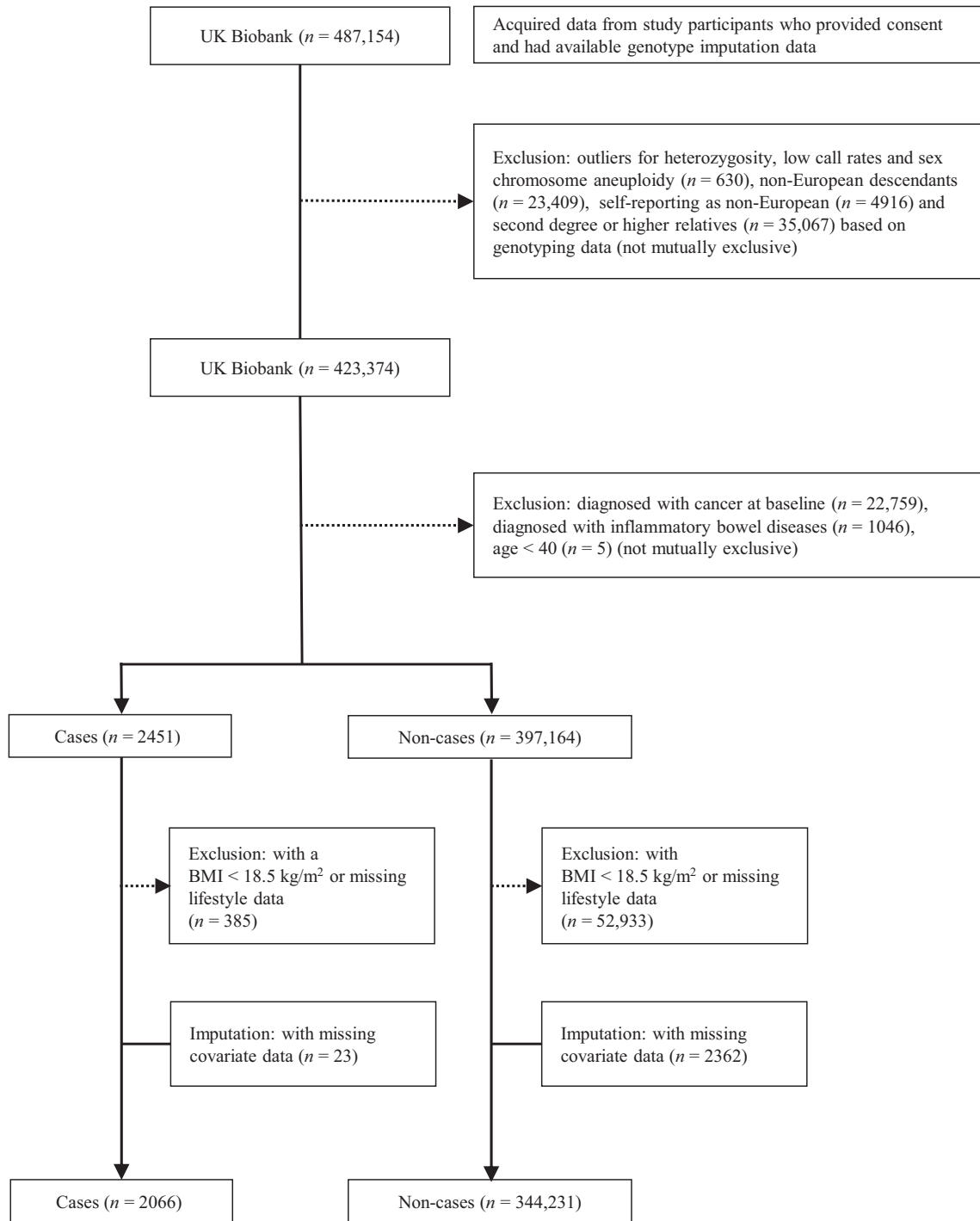


FIGURE 1 Flow diagram of the study population selection.

The associations between the HLS and risk of CRC by PRS categories are shown in [Table 3](#). The HLS was significantly associated with a reduced risk of CRC, regardless of PRS categories, for all participants combined and for men in both HLS1 and HLS2 analyses. For women, reduced risk associated with a healthy HLS score was also observed, although the association was only statistically significant for HLS2 in the high PRS category. The association of the HLS with CRC risk

was stronger among those with a high PRS than those with a low or intermediate PRS, and tests for additive interactions were statistically significant (HLS1: $P = 0.02$ for all participants combined and $P = 0.04$ for men; HLS2: $P = 0.00005$ for all participants combined, $P = 0.002$ for men, and $P = 0.01$ for women). However, there were no significant multiplicative interactions of either HLS1 or HLS2 with PRS. This pattern of the association was more apparent in those who reported not having

TABLE 1 Comparisons of baseline characteristics by colorectal cancer cases and noncases in the UK Biobank¹

Characteristics	All participants			P value
	combined (n = 346,297)	Cases (n = 2066)	Noncases (n = 344,231)	
Age at enrollment, y	58 (13)	62 (9)	58 (13)	<0.0001
Male sex	161,736 (46.7)	1194 (57.8)	160,542 (46.6)	<0.0001
Family history of colorectal cancer	38,363 (11.1)	306 (14.8)	38,057 (11.1)	<0.0001
Education				<0.0001
College or University degree	120,494 (34.8)	628 (30.4)	119,866 (34.8)	
Some professional qualifications	96,818 (28.0)	555 (26.9)	96,263 (28.0)	
Secondary education	77,868 (22.5)	478 (23.1)	77,390 (22.5)	
None of the above	49,143 (14.2)	384 (18.6)	48,759 (14.2)	
Missing	1974 (0.6)	21 (1.0)	1953 (0.6)	
Living in Townsend least deprived area	58,298 (16.8)	350 (16.9)	57,948 (16.8)	0.27
Ever received bowel screening ²	106,602 (30.8)	740 (35.8)	105,862 (30.8)	<0.0001
PRS ³				<0.0001
Low	115,157 (33.3)	415 (20.1)	114,742 (33.3)	
Intermediate	115,393 (33.3)	645 (31.2)	114,748 (33.3)	
High	115,747 (33.4)	1006 (48.7)	114,741 (33.3)	

¹Values are medians (IQRs) or numbers (%) unless otherwise indicated. HLS, healthy lifestyle score; PRS, polygenic risk score.

²Bowel screening includes tests for blood in the stool/feces or a colonoscopy or a sigmoidoscopy.

³The PRS was categorized as low (≤ 7.6613), intermediate (7.6614–8.0567), and high (> 8.0567) based on the tertile distribution of PRS in noncases.

received any bowel screening than those who reported having received a screening before baseline. The number of cases and noncases for each group is presented in **Supplemental Table 4**. In addition, the associations between the HLS and risk of CRC by a family history of CRC are shown in **Supplemental Table 5**. An inverse association of the HLS with risk of CRC was found regardless of family history of CRC, although not all associations were statistically significant. The association of the HLS with CRC risk was stronger among those without a family history of CRC than those with a family history of CRC. However, there were no significant additive or multiplicative interactions of either HLS1 or HLS2 with a family history of CRC.

We calculated the cumulative risk of CRC from age 40 to 75 y for groups defined jointly by HLS and PRS (**Figure 2** and **Supplemental Table 6**). In the HLS1 analysis (**Figure 2A**), compared with those with a low PRS and a healthy HLS (risk = 2.09%), individuals with a low PRS and an unhealthy HLS had a 2.77% risk of CRC (an excess risk of 0.68% due to HLS alone) and individuals with a high PRS and healthy HLS had a 4.52% risk of CRC (an excess risk of 2.43% due to PRS alone). However, among those with a high PRS and unhealthy HLS, the risk was increased to 6.40%, an excess risk of 4.31% compared with the low PRS/healthy HLS group, including an excess risk of 1.20% due to an interaction of PRS and HLS. This pattern of associations was more evident in HLS2 (**Figure 2B**). These associations were also supported with results from analyses using RERI, in which a significant RERI was found in groups with a high PRS/unhealthy HLS (RERI_{HLS1}: 0.58; 95% CI: 0.06, 1.10; RERI_{HLS2}: 1.09; 95% CI: 0.60, 1.59). Again, this pattern of associations was more evident in men than in women (**Table 3**, **Supplemental Figures 1** and **2**) and those who reported not having received any bowel screening than those who reported having received a screening before baseline (**Table 3**, **Figure 3**, and **Supplemental Figure 3**). However, even

for those who reported having received a bowel screening, there were significant interactions of PRS with HLS2 (**Table 3** and **Supplemental Figure 3**). Also, graphs for cumulative risks by age are presented to show how cumulative risk may vary across risk groups defined by PRS and HLS (**Supplemental Figure 4**).

Furthermore, we assessed the joint association of the HLS and PRS from a multiple Cox regression model with the reference of a low PRS/healthy HLS according to sex (**Supplemental Figure 5**). There was significant heterogeneity between men and women (P for heterogeneity = 0.001 for HLS1 and 0.026 for HLS2).

Discussion

Using data from a large population-based cohort study, we found that adherence to a healthy lifestyle was associated with a significantly reduced risk of developing CRC, regardless of the risk determined by genetic factors. Intriguingly, the reduction in CRC risk was most evident among those at the highest genetic risk, following a greater than additive interaction model. Our study provides strong evidence to support lifestyle modifications for CRC primary prevention.

Several previous studies have evaluated potential gene–environment interactions in CRC risk. For example, several other studies showed that dietary factors may interact with genetic variants to modify the risk of CRC (28–32). Figueiredo et al. reported a significant interaction between processed meat consumption and the variant rs4143094 (28). Gong et al. showed an interaction between alcohol consumption and variants in the 9q22.32/HIATL1 region (29). However, very few studies have quantified potential interactions of aggregated genetic susceptibility and overall lifestyle in CRC risk. In a study by Carr et al. using an HLS derived from 5 modifiable lifestyle factors (smoking, alcohol consumption, diet, physical activity,

TABLE 2 Associations between components of the healthy lifestyle score and risk of colorectal cancer; results from the UK Biobank^{1,2}

HLS components	All participants combined		Men		Women	
	Cases	HR (95% CI)	Cases	HR (95% CI)	Cases	HR (95% CI)
BMI, kg/m ²						
18.5–24.9	577	1 (Ref)	253	1 (Ref)	324	1 (Ref)
25.0–29.9	957	1.11 (0.999, 1.23)	593	1.13 (0.98, 1.31)	364	1.13 (0.97, 1.32)
≥30.0	532	1.17 (1.04, 1.32)	348	1.34 (1.13, 1.57)	184	0.98 (0.82, 1.18)
Waist-to-hip ratio (by sex-specific tertiles) ³						
T ₁ (low)	503	1 (Ref)	271	1 (Ref)	232	1 (Ref)
T ₂	699	1.24 (1.11, 1.40)	406	1.34 (1.15, 1.56)	293	1.13 (0.95, 1.35)
T ₃	864	1.41 (1.26, 1.58)	517	1.53 (1.32, 1.78)	347	1.26 (1.06, 1.49)
Physical activity 10+ min, d/wk						
Moderate 6–7 and vigorous 3–5 or 6–7	287	1 (Ref)	172	1 (Ref)	115	1 (Ref)
Moderate 1–5 or vigorous 1–2	1545	1.07 (0.95, 1.22)	888	1.14 (0.97, 1.34)	657	0.98 (0.80, 1.19)
None	234	1.09 (0.92, 1.30)	134	1.12 (0.89, 1.40)	100	1.04 (0.80, 1.36)
Sedentary time, h/d, by tertiles						
<3	427	1 (Ref)	208	1 (Ref)	219	1 (Ref)
3–4.9	802	1.08 (0.96, 1.21)	446	1.09 (0.92, 1.28)	356	1.09 (0.92, 1.29)
≥5	837	1.09 (0.96, 1.22)	540	1.11 (0.95, 1.31)	297	1.07 (0.89, 1.28)
Consumption of processed and red meat						
<4 times per wk	214	1 (Ref.)	67	1 (Ref)	147	1 (Ref)
5–7 times per wk	1375	1.17 (1.01, 1.35)	758	1.40 (1.09, 1.79)	617	1.09 (0.91, 1.31)
≥8 times per wk	477	1.28 (1.09, 1.51)	369	1.69 (1.30, 2.19)	108	0.92 (0.72, 1.18)
Vegetable and fruit intake						
>5 servings per d	620	1 (Ref)	306	1 (Ref)	314	1 (Ref)
2–4 servings per d	1185	0.99 (0.90, 1.09)	694	0.98 (0.86, 1.12)	491	0.999 (0.87, 1.15)
≤1 serving per d	261	1.07 (0.93, 1.24)	194	1.09 (0.91, 1.30)	67	1.03 (0.79, 1.34)
Alcohol consumption						
Nonuser or seldom ⁴	519	1 (Ref)	195	1 (Ref)	324	1 (Ref)
1–4 times per wk	988	1.003 (0.90, 1.12)	610	1.19 (1.01, 1.39)	378	0.86 (0.74, 1.001)
Almost daily	559	1.16 (1.02, 1.31)	389	1.36 (1.14, 1.62)	170	0.98 (0.81, 1.19)
Tobacco smoking						
Never	985	1 (Ref)	485	1 (Ref)	500	1 (Ref)
Former and current, light ⁵	960	1.18 (1.08, 1.30)	632	1.27 (1.13, 1.43)	328	1.06 (0.92, 1.22)
Current, heavy	121	1.34 (1.11, 1.63)	77	1.35 (1.06, 1.72)	44	1.36 (0.99, 1.86)
HLS1 (met/not met guideline score) ⁶						
Unhealthy (≤1)	783	1 (Ref)	626	1 (Ref)	157	1 (Ref)
Intermediate (2, 3)	906	0.83 (0.75, 0.91)	466	0.81 (0.72, 0.91)	440	0.93 (0.77, 1.11)
Healthy (≥4, healthy)	377	0.71 (0.63, 0.81)	102	0.60 (0.49, 0.74)	275	0.84 (0.69, 1.02)
<i>P</i> for trend		<0.0001		<0.0001		0.07
HLS2 (sex-specific weighted score) ⁷						
Unhealthy	920	1 (Ref)	555	1 (Ref)	365	1 (Ref)
Intermediate	654	0.78 (0.70, 0.86)	378	0.75 (0.66, 0.86)	276	0.81 (0.70, 0.95)
Healthy	492	0.65 (0.58, 0.73)	261	0.59 (0.51, 0.69)	231	0.74 (0.62, 0.87)
<i>P</i> for trend		<0.0001		<0.0001		0.0003

¹HLS, healthy lifestyle score.

²Adjusted for age at enrollment, sex (only in analysis of all participants combined), education, Townsend deprivation index, bowel screening, and family history. Stratified by birth cohort.

³For the HLS1 component, the waist-to-hip ratio was categorized according to WHO guidelines. For the HLS2 component, the waist-to-hip ratio was categorized into tertiles T₁ (<0.91), T₂ (0.91–0.96) and T₃ (>0.96) for men and T₁ (<0.78), T₂ (0.78–0.84) and T₃ (>0.84) for women, according to the distribution of waist-to-hip ratio in noncases.

⁴Seldom: special occasions only, and 1–3 times a month.

⁵Light current smoking is defined as smoking <20 pack-y. Heavy current smoking is defined as smoking ≥20 pack-y.

⁶We provide a detailed explanation for each category in Supplementary Table 1. Each category was scored 0 or 1, with 1 representing having met the guideline. The HLS1 was then constructed by taking the sum of the scores. The HLS1 ranged from 0 to 8, and was analyzed in this study according to 3 categories: unhealthy (score 0, 1), intermediate (2, 3), and healthy (≥4), based on the tertile distribution of HLS1 in noncases.

⁷Each category was scored according to the estimated sex-specific weights, log_e-HR, for the 8 components to construct the HLS2 from a multivariable logistic regression analysis. The HLS2 was analyzed in this study according to 3 categories: unhealthy (>1.39), intermediate (1.04–1.39), and healthy (<1.04) for men and unhealthy (>0.28), intermediate (0.12–0.28), and healthy (<0.12) for women, based on the tertile distribution of HLS2 in noncases.

and body fatness) and genetic risk scores based on 53 GWAS-identified CRC risk SNPs, the risk of CRC was found to decrease according to healthy lifestyle factors, regardless of genetic risk (14). However, similar to our study, no significant multiplicative

interaction was found between genetic and lifestyle factors. Additive interactions, however, were not evaluated in that study. In a pooled analysis including both cohort and case-control studies, Wang et al. reported a statistically significant additive

TABLE 3 Association between the healthy lifestyle score and risk of colorectal cancer stratified by polygenic risk scores; results from the UK Biobank¹

PRS category ²	HLS category						<i>P</i> value ³
	Unhealthy			Healthy			
	Cases	HR (95% CI)	Cases	HR (95% CI)	Cases	HR (95% CI)	
All participants combined							
Low	159	1.00 (Ref)	175	0.79 (0.63, 0.98)	81	0.76 (0.57, 1.01)	0.04
Intermediate	244	1.00 (Ref)	282	0.83 (0.70, 0.995)	119	0.73 (0.58, 0.92)	0.005
High	380	1.00 (Ref)	449	0.85 (0.74, 0.98)	177	0.71 (0.58, 0.85)	0.0003
Men							<i>P</i> _{additive interaction} 0.02
Low	126	1.00 (Ref)	91	0.79 (0.61, 1.04)	20	0.59 (0.37, 0.95)	<i>P</i> _{multiplicative interaction} 0.80
Intermediate	190	1.00 (Ref)	141	0.81 (0.65, 1.01)	34	0.67 (0.47, 0.97)	0.01
High	310	1.00 (Ref)	234	0.82 (0.69, 0.97)	48	0.58 (0.43, 0.79)	0.0001
Women							<i>P</i> _{additive interaction} 0.04
Low	33	1.00 (Ref)	84	0.83 (0.56, 1.25)	61	0.90 (0.59, 1.38)	<i>P</i> _{multiplicative interaction} 0.87
Intermediate	54	1.00 (Ref)	141	0.90 (0.66, 1.24)	85	0.80 (0.56, 1.13)	0.77
High	70	1.00 (Ref)	215	1.01 (0.77, 1.32)	129	0.88 (0.65, 1.18)	0.19
Never screened ⁵							<i>P</i> _{additive interaction} 0.49
Low	88	1.00 (Ref)	103	0.78 (0.58, 1.05)	49	0.76 (0.52, 1.11)	<i>P</i> _{multiplicative interaction} 0.96
Intermediate	160	1.00 (Ref)	173	0.75 (0.60, 0.94)	70	0.60 (0.45, 0.81)	0.12
High	249	1.00 (Ref)	302	0.83 (0.69, 0.98)	110	0.61 (0.48, 0.78)	0.0005
Ever screened ⁵							<i>P</i> _{additive interaction} 0.003
Low	70	1.00 (Ref)	70	0.77 (0.55, 1.09)	32	0.76 (0.49, 1.19)	<i>P</i> _{multiplicative interaction} 0.62
Intermediate	83	1.00 (Ref)	103	0.95 (0.71, 1.28)	48	0.98 (0.67, 1.43)	0.17
High	125	1.00 (Ref)	143	0.90 (0.71, 1.16)	66	0.93 (0.67, 1.28)	0.87
All participants combined							0.56
Low	179	1.00 (Ref)	132	0.82 (0.66, 1.03)	104	0.73 (0.57, 0.93)	<i>P</i> _{additive interaction} 0.99
Intermediate	275	1.00 (Ref)	211	0.85 (0.71, 1.01)	159	0.72 (0.59, 0.88)	<i>P</i> _{multiplicative interaction} 0.69
High	466	1.00 (Ref)	311	0.72 (0.63, 0.84)	229	0.59 (0.50, 0.69)	0.01
Men							<i>P</i> _{additive interaction} 0.00005
Low	113	1.00 (Ref)	70	0.70 (0.52, 0.95)	54	0.63 (0.45, 0.87)	<i>P</i> _{multiplicative interaction} 0.18
Intermediate	158	1.00 (Ref)	125	0.87 (0.69, 1.10)	82	0.66 (0.50, 0.86)	0.003
High	284	1.00 (Ref)	183	0.71 (0.59, 0.86)	125	0.55 (0.44, 0.68)	0.0001
							<i>P</i> _{additive interaction} 0.002
							<i>P</i> _{multiplicative interaction} 0.63

(Continued)

TABLE 3 (Continued)

PRS category ²	HLS category						P value ³
	Unhealthy		Intermediate		Healthy		
	Cases	HR (95% CI)	Cases	HR (95% CI)	Cases	HR (95% CI)	
Women							
Low	66	1.00 (Ref)	62	1.03 (0.72, 1.46)	50	0.90 (0.62, 1.31)	0.61
Intermediate	117	1.00 (Ref)	86	0.81 (0.61, 1.08)	77	0.81 (0.60, 1.09)	0.14
High	182	1.00 (Ref)	128	0.75 (0.60, 0.94)	104	0.65 (0.51, 0.83)	0.0004
							<i>P</i> _{additive interaction} 0.01
							<i>P</i> _{multiplicative interaction} 0.16
Never screened ⁵							
Low	107	1.00 (Ref)	80	0.80 (0.60, 1.07)	53	0.57 (0.41, 0.80)	0.001
Intermediate	176	1.00 (Ref)	126	0.77 (0.61, 0.97)	101	0.67 (0.52, 0.86)	0.001
High	311	1.00 (Ref)	191	0.64 (0.53, 0.76)	159	0.57 (0.47, 0.69)	<0.0001
							<i>P</i> _{additive interaction} 0.0006
							<i>P</i> _{multiplicative interaction} 0.62
Ever screened ⁵							
Low	70	1.00 (Ref)	51	0.85 (0.59, 1.22)	51	1.02 (0.71, 1.48)	0.98
Intermediate	96	1.00 (Ref)	82	0.98 (0.73, 1.32)	56	0.80 (0.58, 1.12)	0.22
High	149	1.00 (Ref)	118	0.93 (0.73, 1.18)	67	0.63 (0.47, 0.84)	0.002
							<i>P</i> _{additive interaction} 0.01
							<i>P</i> _{multiplicative interaction} 0.051

¹HRs were adjusted for age at enrollment, sex (only for all participants combined), education, Townsend deprivation index, bowel screening (only for all participants combined, men, and women), family history, top 5 principal components for ancestry, and genotyping batch. Stratified by birth cohort. HLS, healthy lifestyle score; PRS, polygenic risk score.

²The PRS was categorized as low (score: ≤ 7.6613), intermediate (7.6614–8.0567), and high (> 8.0567) based on the tertile distribution of PRS in noncases.

³*P* for trend unless otherwise indicated.

⁴Each category was scored 0 or 1, with 1 representing having met the guideline. The HLS1 was then constructed by taking the sum of the scores. The HLS1 ranged from 0 to 8, and was analyzed in this study according to 3 categories: unhealthy (score 0, 1), intermediate (2, 3), and healthy (≥ 4), based on the tertile distribution of HLS1 in noncases.

⁵Bowel screening includes tests for blood in the stool/feces or a colonoscopy or a sigmoidoscopy (missing $n = 4737$). Cohort members who reported not having received any bowel screening before the baseline survey were placed into the “never screened” group, and those who reported having received a screening test for CRC before the baseline survey were categorized as the “ever screened” group.

⁶Each category was scored according to the estimated sex-specific weights, log_e-HR, for the 8 components to construct the HLS2 from a multivariable logistic regression analysis. The HLS2 was analyzed in this study according to 3 categories: unhealthy (> 1.39), intermediate (1.04–1.39), and healthy (< 1.04) for men and unhealthy (> 0.28), intermediate (0.12–0.28), and healthy (< 0.12) for women, based on the tertile distribution of HLS2 in noncases.

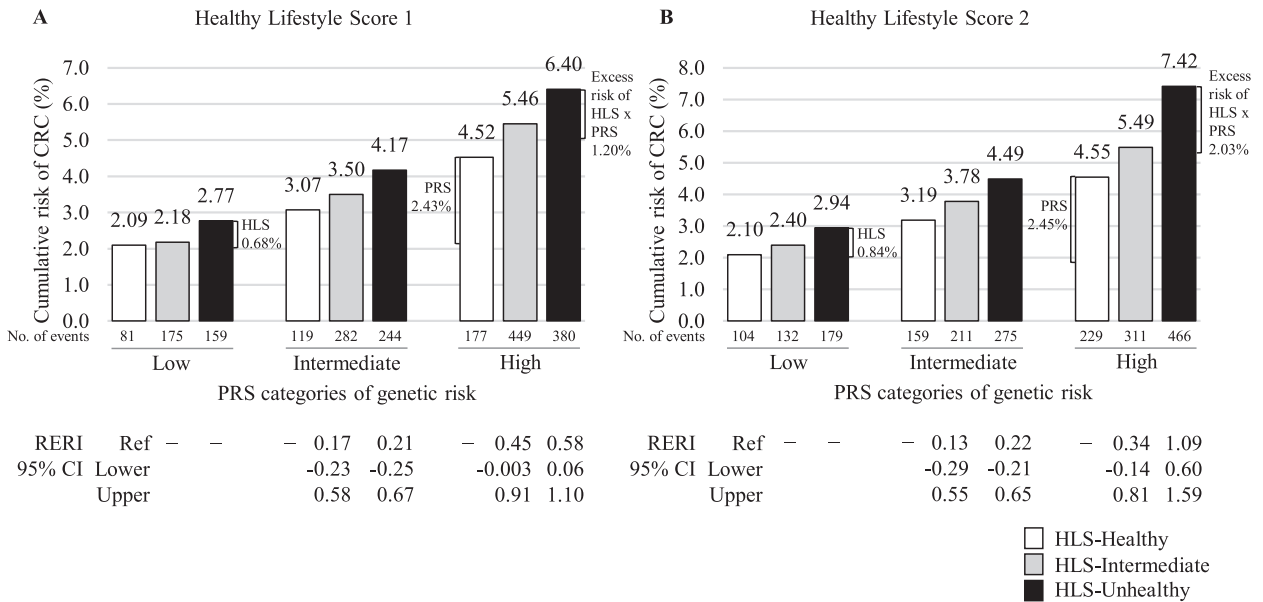


FIGURE 2 Estimates of cumulative risk of developing CRC from age 40 to 75 y defined by HLS [HLS1 (A) and HLS2 (B)] and PRS among all cohort members in the UK Biobank. Cumulative risk was adjusted for age at enrollment, sex, education, Townsend deprivation index, bowel screening, family history, top 5 principal components for ancestry, genotyping batch, and birth cohort. To estimate the RERI, the healthiest HLS and the lowest genetic risk (low PRS) groups were used as the reference. CRC, colorectal cancer; HLS, healthy lifestyle score; PRS, polygenic risk score; RERI, relative excess risk due to interaction.

interaction between modifiable factors and nonmodifiable risk scores derived using information on height, CRC family history, and a genetic risk score combining the estimated effects of 63 GWAS-identified SNPs (15). In our study, we included additional genetic risk variants to measure genetic susceptibility and used prospective cohort study design, which reduces the possibility of

recall bias to improve the validity of study results. Nevertheless, both studies provided evidence for additive interactions between lifestyle and genetic factors in CRC risk. Furthermore, in our study, the additive interactions of PRS and HLS were more apparent among those who reported not having received any bowel screening than those who reported having received a bowel

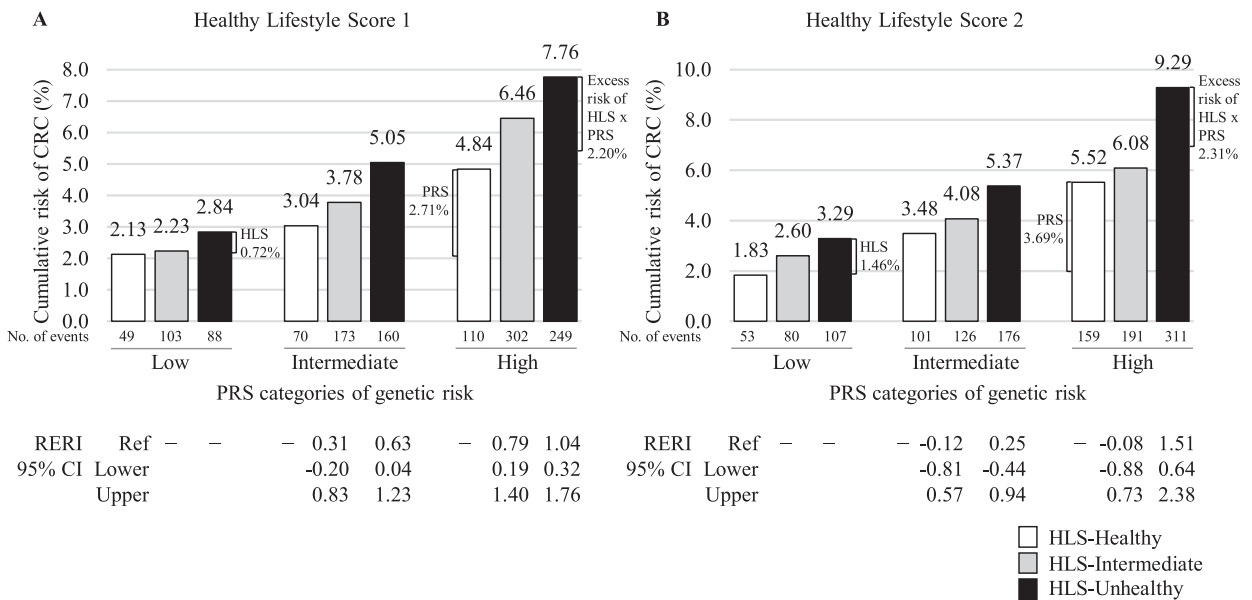


FIGURE 3 Estimates of cumulative risk of developing colorectal cancer from age 40 to 75 y defined by the HLS [HLS1 (A) and HLS2 (B)] and PRS among cohort members who reported not having received any bowel screening in the UK Biobank. Cumulative risk was adjusted for age at enrollment, sex, education, Townsend deprivation index, family history, top 5 principal components for ancestry, genotyping batch, and birth cohort. To estimate the RERI, the healthiest HLS and the lowest genetic risk (low PRS) groups were used as the reference. CRC, colorectal cancer; HLS, healthy lifestyle score; PRS, polygenic risk score; RERI, relative excess risk due to interaction.

screening before baseline. This finding is supported by a recent case-control study showing that the associations of CRC risk with healthy lifestyle and genetic factors were modified by CRC screening (33). In addition, we analyzed the associations between the HLS and risk of CRC by a family history of CRC. However, there were no significant additive or multiplicative interactions between HLS and a family history of CRC. It is possible that shared environmental and genetic factors might affect the family history of cancer. In this study, we adjusted for the family history of CRC in other analyses.

Although we found a similar association pattern between HLS and the risk of CRC between men and women, the association appears to be stronger in men. In men, a high HLS was associated with 30–50% reduction of CRC risk across all PRS categories, whereas no significant association was found for HLS in women except in the high PRS in HLS2. The number of CRC cases in women is smaller than in men, which may partially explain a smaller number of significant findings in women. In addition, previous studies have reported that the association of obesity and alcohol with risk of CRC differed by sex (34–37). In men, obesity was associated with an increased risk of CRC (34). On the other hand, large prospective studies have shown that BMI was associated with a 2-fold increased CRC risk among premenopausal women or young women, but was not associated with CRC risk, or even associated with a small reduction in CRC risk among postmenopausal women or older women (38, 39). In our study, the UK Biobank recruited participants >40 y old, with a mean age of 56, and thus it is likely that more postmenopausal than premenopausal women were investigated in our study. It has also been reported that adult weight gain in men and early obesity in women are more significantly associated with CRC risk (35). Consistent with our results, an earlier study reported that alcohol consumption was associated with an increased CRC risk only in men (36). Although the biologic mechanism of differential association of sex remains unclear, studies have suggested that it might be caused by sex hormones. In postmenopausal women, estrogen is mainly produced in fat tissue, which is the main source of endogenous estrogen after menopause (40). A high endogenous estrogen derived from fat tissue may have a protective effect against CRC risk in postmenopausal women (37, 38, 41). Conversely, in premenopausal women, estrogen derived from fat tissue is insignificant compared with estrogen derived from the ovaries (40), which would not offset the deleterious effects of obesity on CRC risk (38). Other studies have reported that alcohol ingestion may lead to significant and sustained elevations in endogenous estrogen levels (42, 43).

The most notable finding of this study is that there was a significant additive interaction between genetic and lifestyle factors, and the risk reduction by lifestyle modification was particularly evident among men with high genetic risk. Most previous studies used the multiplicative model to evaluate gene–environment interactions and largely ignored additive interactions. Although uncovering multiplicative interactions could lead to significant insights into potential synergistic effects of 2 risk factors, studying additive interactions could provide more intuitive information for disease prevention. By quantifying directly the level of risk reduction through studying additive interactions, we can identify groups of individuals who are more likely to benefit from lifestyle modifications (44). Our study

includes the additive interaction to better understand the interplay between genetics and the environment in cancer prevention.

There are also some limitations to this study. First, because the self-reported questionnaire was used for the HLS, some misclassification errors in exposure assessment may exist. However, this misclassification is likely to be random, which tends to attenuate the association. In other words, the level of risk reduction presented in our study is likely to be underestimated. Second, in our study, the HLS did not include all components of the ACS guidelines and the sedentary time variable did not capture all sedentary activities as the UK Biobank did not collect sufficient data on weight change, whole grain, highly processed foods, refined grain products, sugar-sweetened beverages, and other sedentary activities. It is possible that the influence of healthy lifestyle on CRC risk was underestimated in our study. Future studies including additional components of the ACS guidelines are needed to fully capture the effect of lifestyle on CRC risk. Third, as those CRC-related SNPs were selected from people of European descent, the generalizability of the study findings to other populations should be evaluated in future studies. Strengths of this study include its large sample size through a large prospective cohort study, inclusion of a comprehensive list of GWAS-identified SNPs for CRC risk, and the use of standardized healthy lifestyle behaviors, according to the ACS guidelines.

In conclusion, although the observational nature of the study design precludes proof of causality, our study suggests that healthy lifestyles may reduce the risk of CRC, particularly among individuals with a genetic susceptibility. The benefit of lowering the risk of CRC by adherence to a healthy lifestyle is expected to be the greatest among individuals with highest genetic susceptibility, which can help establish personalized preventive strategies for cancer prevention.

This research was conducted using the UK Biobank Resource under application number 40685. Deidentified data from the UK Biobank were used in this project. The UK Biobank is an open access resource open to bona fide scientists undertaking health-related research that is in the public good. Approved scientists can apply to use this resource.

The authors' responsibilities were as follows—JC, GJ, WZ: designed the research, had full access to all study data, and conducted the research; JC, GJ, WW, WZ: performed the statistical analyses and analyzed the data; JC, WZ: wrote the manuscript and had primary responsibility for final content; all authors: contributed to data interpretation and critical revision of the manuscript; and all authors: read and approved the final manuscript. Author disclosures: The authors report no conflicts of interest.

Data Availability

Data used in this project, along with code books and codes, can be obtained directly from the UK Biobank by submitting a data request proposal.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68(6):394–424.
2. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50(9):1219–24.

3. Jia G, Lu Y, Wen W, Long J, Liu Y, Tao R, Li B, Denny JC, Shu X-O, Zheng W. Evaluating the utility of polygenic risk scores in identifying high-risk individuals for eight common cancers. *JNCI Cancer Spectr* 2020;4(3):pkaa021.
4. Lu Y, Kweon S-S, Tanikawa C, Jia W-H, Xiang Y-B, Cai Q, Zeng C, Schmit SL, Shin A, Matsuo K, et al. Large-scale genome-wide association study of East Asians identifies loci associated with risk for colorectal cancer. *Gastroenterology* 2019;156(5):1455–66.
5. Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, Conti DV, Qu C, Jeon J, Edlund CK, et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet* 2019;51(1):76–87.
6. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet North Am Ed* 2008;371(9612):569–78.
7. Bassett JK, Severi G, English DR, Baglietto L, Krishnan K, Hopper JL, Giles GG. Body size, weight change, and risk of colon cancer. *Cancer Epidemiol Biomarkers Prev* 2010;19(11):2978–86.
8. Larsson SC, Wolk A. Meat consumption and risk of colorectal cancer: a meta-analysis of prospective studies. *Int J Cancer* 2006;119(11):2657–64.
9. Bouvard V, Loomis D, Guyton KZ, Grosse Y, Ghissassi FE, Benbrahim-Tallaa L, Guha N, Mattock H, Straif K. Carcinogenicity of consumption of red and processed meat. *Lancet Oncol* 2015;16(16):1599–600.
10. Bagnardi V, Rota M, Botteri E, Tramacere I, Islami F, Fedirko V, Scotti L, Jenab M, Turati F, Pasquali E, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose–response meta-analysis. *Br J Cancer* 2015;112(3):580–93.
11. Liang PS, Chen T-Y, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. *Int J Cancer* 2009;124(10):2406–15.
12. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. *Int J Cancer* 2009;125(1):171–80.
13. Thomson CA, McCullough ML, Wertheim BC, Chlebowski RT, Martinez ME, Stefanick ML, Rohan TE, Manson JE, Tindle HA, Ockene J, et al. Nutrition and physical activity cancer prevention guidelines, cancer risk, and mortality in the women’s health initiative. *Cancer Prev Res* 2014;7(1):42–53.
14. Carr PR, Weigl K, Jansen L, Walter V, Erben V, Chang-Claude J, Brenner H, Hoffmeister M. Healthy lifestyle factors associated with lower risk of colorectal cancer irrespective of genetic risk. *Gastroenterology* 2018;155(6):1805–15.e5.
15. Wang X, O’Connell K, Jeon J, Song M, Hunter D, Hoffmeister M, Lin Y, Berndt S, Brenner H, Chan AT, et al. Combined effect of modifiable and non-modifiable risk factors for colorectal cancer risk in a pooled analysis of 11 population-based studies. *BMJ Open Gastroenterol* 2019;6(1):e000339.
16. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, 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* 2015;12(3):e1001779.
17. Rock CL, Thomson C, Gansler T, Gapstur SM, McCullough ML, Patel AV, Andrews KS, Bandera EV, Spees CK, Robien K, et al. American Cancer Society guideline for diet and physical activity for cancer prevention. *CA A Cancer J Clin* 2020;70(4):245–71.
18. American Cancer Society. Stay away from tobacco. [Internet]. Atlanta, Georgia: American Cancer Society, Inc.; 2020 [cited 13 July, 2020]. Available from: <https://www.cancer.org/healthy/stay-away-from-tobacco.html>.
19. World Cancer Research Fund. Resources and toolkits. Diet, nutrition, physical activity and colorectal cancer. [Internet]. London, UK: World Cancer Research Fund International; 2017 [cited 13 July, 2020]. Available from: <https://www.wcrf.org/dietandcancer/resources-and-toolkit>.
20. World Health Organization. Waist circumference and waist-hip ratio: Report of a WHO Expert Consultation, Geneva, 8–11 December 2008. Geneva: World Health Organization; 2011.
21. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, George SM, Olson RD. The physical activity guidelines for Americans. *JAMA* 2018;320(19):2020–28.
22. Bradbury KE, Young HJ, Guo W, Key TJ. Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. *J Nutr Sci* 2018;7:e6.
23. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global reference for human genetic variation. *Nature* 2015;526(7571):68–74.
24. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. *Bioinformatics* 2010;26(22):2867–73.
25. Schmit SL, Edlund CK, Schumacher FR, Gong J, Harrison TA, Huyghe JR, Qu C, Melas M, Van Den Berg DJ, Wang H, et al. Novel common genetic susceptibility loci for colorectal cancer. *J Natl Cancer Inst* 2019;111(2):146–57.
26. Li R, Chambless L. Test for additive interaction in proportional hazards models. *Ann Epidemiol* 2007;17(3):227–36.
27. Arthur RS, Wang T, Xue X, Kamensky V, Rohan TE. Genetic factors, adherence to healthy lifestyle behavior, and risk of invasive breast cancer among women in the UK Biobank. *J Natl Cancer Inst* 2020;112(9):893–901.
28. Figueiredo JC, Hsu L, Hutter CM, Lin Y, Campbell PT, Baron JA, Berndt SI, Jiao S, Casey G, Fortini B, et al. Genome-wide diet-gene interaction analyses for risk of colorectal cancer. *PLoS Genet* 2014;10(4):e1004228.
29. Gong J, Hutter CM, Newcomb PA, Ulrich CM, Bien SA, Campbell PT, Baron JA, Berndt SI, Bezieau S, Brenner H, et al. Genome-wide interaction analyses between genetic variants and alcohol consumption and smoking for risk of colorectal cancer. *PLoS Genet* 2016;12(10):e1006296.
30. Liu AY, Scherer D, Poole E, Potter JD, Curtin K, Makar K, Slattery ML, Caan BJ, Ulrich CM. Gene-diet-interactions in folate-mediated one-carbon metabolism modify colon cancer risk. *Mol Nutr Food Res* 2013;57(4):721–34.
31. Cotterchio M, Boucher BA, Manno M, Gallinger S, Okey AB, Harper PA. Red meat intake, doneness, polymorphisms in genes that encode carcinogen-metabolizing enzymes, and colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev* 2008;17(11):3098–107.
32. Turner F, Smith G, Sachse C, Lightfoot T, Garner RC, Wolf CR, Forman D, Bishop DT, Barrett JH. Vegetable, fruit and meat consumption and potential risk modifying genes in relation to colorectal cancer. *Int J Cancer* 2004;112(2):259–64.
33. Carr PR, Weigl K, Edelmann D, Jansen L, Chang-Claude J, Brenner H, Hoffmeister M. Estimation of absolute risk of colorectal cancer based on healthy lifestyle, genetic risk, and colonoscopy status in a population-based study. *Gastroenterology* 2020;159(1):129–38.e9.
34. Larsson SC, Wolk A. Obesity and colon and rectal cancer risk: a meta-analysis of prospective studies. *Am J Clin Nutr* 2007;86(3):556–65.
35. Kim H, Giovannucci EL. Sex differences in the association of obesity and colorectal cancer risk. *Cancer Causes Control* 2017;28(1):1–4.
36. Wu A, Paganini-Hill A, Ross R, Henderson B. Alcohol, physical activity and other risk factors for colorectal cancer: a prospective study. *Br J Cancer* 1987;55(6):687–94.
37. Tabung FK, Liu L, Wang W, Fung TT, Wu K, Smith-Warner SA, Cao Y, Hu FB, Ogino S, Fuchs CS, et al. Association of dietary inflammatory potential with colorectal cancer risk in men and women. *JAMA Oncol* 2018;4(3):366–73.
38. Terry PD, Miller AB, Rohan TE. Obesity and colorectal cancer risk in women. *Gut* 2002;51(2):191–4.
39. Terry P, Giovannucci E, Bergkvist L, Holmberg L, Wolk A. Body weight and colorectal cancer risk in a cohort of Swedish women: relation varies by age and cancer site. *Br J Cancer* 2001;85(3):346–9.
40. Bernstein L, Ross RK. Endogenous hormones and breast cancer risk. *Epidemiol Rev* 1993;15(1):48–65.
41. Murphy N, Strickler HD, Stanczyk FZ, Xue X, Wassertheil-Smoller S, Rohan TE, Ho GYF, Anderson GL, Potter JD, Gunter MJ. A prospective evaluation of endogenous sex hormone levels and colorectal cancer risk in postmenopausal women. *JNCIJ* 2015;107(10):djv210.
42. Ginsburg ES, Mello NK, Mendelson JH, Barbieri RL, Teoh SK, Rothman M, Gao X, Sholar JW. Effects of alcohol ingestion on estrogens in postmenopausal women. *JAMA* 1996;276(21):1747–51.
43. Pöschl G, Seitz HK. Alcohol and cancer. *Alcohol Alcohol* 2004;39(3):155–65.
44. VanderWeele TJ, Knol MJ. A tutorial on interaction. *Epidemiologic Methods* 2014;3(1):33–72.