Published in final edited form as:

Gastroenterology. 2020 July ; 159(1): 241–256.e13. doi:10.1053/j.gastro.2020.03.054.

Risk Factor Profiles Differ for Cancers of Different Regions of the Colorectum

Liang Wang[1,2], **Chun-Han Lo**[2,3], **Xiaosheng He**[4], **Dong Hang**[5], **Molin Wang**[2,6], **Kana Wu**[7], **Andrew T. Chan**[3,7,11,12,13], **Shuji Ogino**[2,9,10,11], **Edward L. Giovannucci**[2,7,8], **Mingyang Song**[2,3,7,12]

[1]Center of Gastrointestinal Surgery, the First Affiliated Hospital, Sun Yat-sen University, Guangzhou, P.R. China

[2]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[3]Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

[4]Department of Colorectal Surgery, the Six Affiliated Hospital, Sun Yat-sen University, Guangzhou, P.R. China

[5]Department of Epidemiology and Biostatistics, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing, P.R. China

[6]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[7]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[8]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

[9]Department of Oncologic Pathology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA

[10]Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

[11]Broad Institute of MIT and Harvard, Cambridge, MA, USA

[12]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Corresponding author: Mingyang Song; Department of Epidemiology, Harvard T.H. Chan School of Public Health, 667 Huntington Avenue, Kresge 906A, Boston, MA 02115; Tel: +01-617-432-1301; Fax: +01-617-566-7805, mingyangsong@mail.harvard.edu.

Author contributions: Study concept and design: LW, MS; acquisition of data: CHL, XH, DH, KW, MLW, ELG, ATC, MS; statistical analysis: LW; interpretation of data: all authors; drafting of the manuscript: LW, MS; critical revision of the manuscript for important intellectual content: CHL, XH, DH, KW, SO, ATC, ELG.

[13]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Abstract

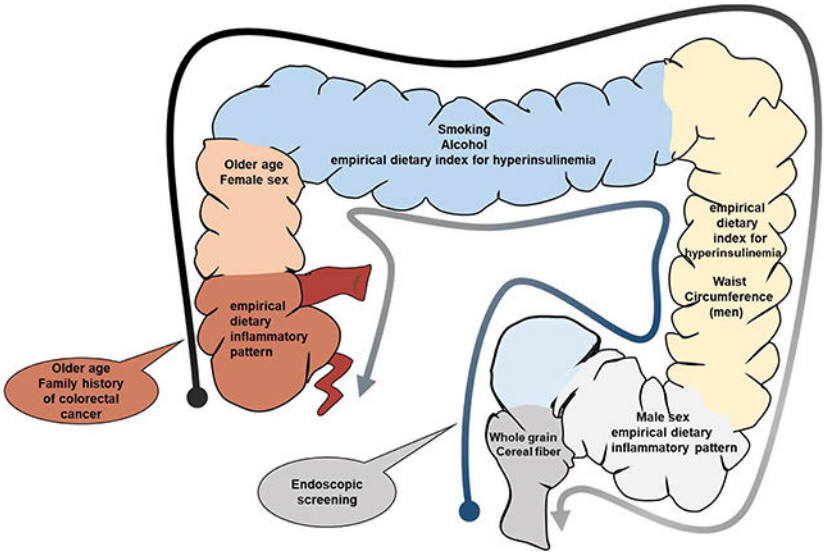
Background & Aims—The molecular features of colorectal tumors differ with their anatomic location. Colorectal tumors are usually classified as proximal or distal. We collected data from 3 cohorts to identify demographic, clinical, anthropometric, lifestyle, and dietary risk factors for colorectal cancer (CRC) at 7 anatomic subsites. We examined whether the associations differ among refined subsites and whether there are trends in associations from cecum to rectum.

Methods—We collected data from the Nurses' Health Study, Nurses' Health Study 2, and Health Professionals Follow-up Study (45,351 men and 178,016 women, followed for a median 23 years) on 24 risk factors in relation to risk of cancer in cecum, ascending colon, transverse colon, descending colon, sigmoid colon, rectosigmoid junction, and rectum. Hazard ratios were estimated using Cox proportional hazards regression. We tested for linear and non-linear trends in associations with CRC among subsites and within proximal colon, distal colon, and rectum.

Results—We documented 3058 cases of CRC (474 in cecum, 633 in ascending colon, 250 in transverse colon, 221 in descending colon, 750 in sigmoid colon, 202 in rectosigmoid junction, and 528 in rectum). The positive associations with cancer risk decreased, from cecum to rectum, for age and family history of CRC. In contrast, the inverse associations with cancer risk increased, from cecum to rectum, for endoscopic screening and intake of whole grains, cereal fiber, and processed red meat. There was a significant non-linear trend in the association between CRC and female sex, with hazards ratios ranging from 1.73 for ascending colon cancer to 0.54 for sigmoid colon cancer. For proximal colon cancers, the association with alcohol consumption and smoking before age 30 years increased from the cecum to transverse colon. For distal colon cancers, the positive association with waist circumference in men was greater for descending vs sigmoid colon cancer.

Conclusions—In an analysis of 3058 cases of CRC, we found that risk factor profiles differed for cancers along the colorectum. Proximal vs distal classifications are not sufficient to encompass the regional variations in colorectal tumor features and risk factors.

Graphical Abstract



Keywords

epidemiology; precision prevention; microbiome; spatial variation

Introduction

Colorectal cancer (CRC) is a heterogeneous disease, consisting of etiologically and clinically different subtypes.[1–3] Substantial evidence indicates that the molecular features and risk factors of CRC vary by anatomic subsites, namely proximal colon, distal colon, and rectum.[4–8] These data have provided important insights into the etiologic heterogeneity of CRC. However, accumulating evidence indicates the heterogeneity even within subsites defined by proximal, distal, and rectal location. There does not appear to be an abrupt change in clinicopathological and molecular features between anatomic boundaries such as the splenic flexure, sigmoid and rectosigmoid.[9–17] These data challenge the oversimplified classification of CRC according to main subsites and highlight the need for investigations of refined CRC subsites.

Previously, we examined major molecular features of CRC, including microsatellite instability (MSI), CpG island methylator phenotype (CIMP), and *BRAF* mutation, according to nine anatomic subsites; and found that the positivity of these markers gradually increased from the rectum to the ascending colon.[13, 14] Another study showed that the prevalence of *TP53, KRAS, BRAF, PIK3CA,* and *CTNNB1* mutations differed by refined anatomic location, even within the main subsites.[11] Moreover, increasing data support the important role of the gut microbiota in CRC. It is well known that the composition and abundance of the gut microbiota vary considerably across anatomic sublocations in the colorectum.[18–21] Together, these data indicate the importance of accounting for CRC subsites for better understanding the etiology of CRC.

Leveraging the rich epidemiologic data in three large US cohorts, we investigated 24 demographic, clinical, anthropometric, lifestyle, and dietary risk factors in relation to CRC

across seven anatomic subsites. We examined whether the associations differed across these specific subsites and whether there was any linear trend in the associations from the cecum to rectum.

Methods

Study population

The Nurses' Health Study (NHS) enrolled 121,700 US registered female nurses aged 30–55 years in 1976. Nurses' Health Study 2 (NHS2) included 116,430 registered US female nurses aged 25 to 42 years at time of study entry in 1989. The Health Professionals Follow-up Study (HPFS) enrolled 51,529 US male health professionals aged 40–75 years in 1986. Details about the three cohorts have been described elsewhere.[22–24] Briefly, participants were mailed a questionnaire inquiring about their medical history and lifestyle factors at baseline, and every two years thereafter. Dietary data were collected and updated every four years using the validated food frequency questionnaires (FFQs) since 1980 in the NHS, 1991 in the NHS2, and 1986 in the HPFS. In the present analysis, we used these years as the study baseline.

At baseline, we excluded participants with a history of inflammatory bowel disease or cancer (except for non-melanoma skin cancer) and those with missing FFQs, a high number of blank items on their FFQs (>70), or with implausible caloric intakes (men: <800 or >4,200 kcal/d; women: <600 or >3,500 kcal/d). This study was approved by the Institutional Review Board at the Brigham and Women's Hospital and the Harvard T.H. Chan School of Public Health, and those of participating registries as required.

Ascertainment of CRC cases

On each biennial follow-up questionnaire, participants were asked whether they were diagnosed with CRC during the previous 2 years. For participants who reported CRC diagnosis, we asked for their permission to acquire medical records and pathologic reports. Study physicians, blinded to exposure data, reviewed all medical records to confirm CRC diagnosis and to record the disease stage, histologic findings, and tumor location. We defined proximal cancers as those that occurred in the cecum, ascending colon, and transverse colon; distal colon cancers as those in the descending and sigmoid colon; and rectal cancers as those in the rectosigmoid junction and rectum. Cancers in the hepatic flexure were classified as transverse colon cancer and those in the splenic flexure as descending colon cancer. We also assessed the distribution of major molecular markers (microsatellite instability (MSI), CpG island methylator phenotype (CIMP), and *BRAF* and *KRAS* mutations) according to tumor subsites (Supplementary Figure 1; detailed`laboratory methods are provided in the Supplementary Methods). [25–27]

Assessment of risk factors

We obtained data for 24 CRC risk factors. Non-dietary exposures included age, sex, family history of CRC in a first-degree relative, smoking history, height, body mass index (BMI), BMI at age 18 for female and age 21 for male, leisure-time physical activity, aspirin use, history of endoscopic screening, and alcohol consumption.[28–30] Details of assessment of

these risk factors are described in the Supplementary Methods. Using the semi-quantitative FFQ data, we assessed intake of total fiber, cereal fiber, whole grains, total red meat, processed red meat, unprocessed red meat, folate, calcium, marine omega-3 fatty acid intake (including eicosapentaenoic acid, docosahexaenoic acid, and docosapentaenoic acid), and vitamin D.[31] Supplement use was included in calculation of total nutrient intake which was further adjusted for total caloric intake by the residual method.[32] Furthermore, to examine the overall dietary patterns, we utilized Empirical Dietary Inflammatory Pattern (EDIP) and Empirical Dietary Index for Hyperinsulinemia (EDIH). These dietary indices have shown robust associations with inflammatory and insulin biomarkers, respectively. Details about the development and validation of the EDIP and EDIH have been previously described.[33–36]

Statistical analysis

Participants were followed until the diagnosis of CRC, death, or the end of the follow up (June 1, 2014 for the NHS and the HPFS, and June 1, 2013 for the NHS2), whichever occurred first. We used the subsite-stratified multivariable Cox proportional hazards regression model to calculate the hazard ratios (HRs) and 95% confidence intervals (CIs) of CRC at each anatomic subsite in relation to risk factors.[37] To test for the overall heterogeneity and potential linear trend in the magnitude of the associations across subsites, we used the meta-regression method with a subtype-specific random effect term by treating CRC subsites as a nominal and ordinal variable, respectively.[37, 38] We also performed a heterogeneity test within the main subsites (i.e., proximal colon, distal colon, and rectum) using the likelihood ratio test, by comparing the model in which the association with exposures was allowed to vary by subsite to a model in which all the associations were held constant.

All Cox models were stratified by age, sex, and calendar time except when age or sex was considered as the main exposure. All models were adjusted for race and the non-dietary risk factors. Details of statistical analysis are described in the Supplementary Methods. Dietary factors were categorized into quintiles within each cohort and HRs were calculated using the median of each quintile as a continuous variable. We calculated the HRs per certain increment for continuous variables. The unit of increment was chosen based on the literature and to reflect the distribution of the studied exposure in the US population. Given the known sex difference, we examined sex-specific associations for BMI and waist circumference.

Given the increasing incidence of young-onset CRC in the US and other countries, we characterized the subsite-specific risk factor profiles of CRC according to age at diagnosis. Due to the small number of cases younger than age 50 years in our cohorts (n=199), we used age of 60 years as the cutoff. For younger-onset CRC (<60 years), the follow-up ended when participants reached age 60 years. For older-onset CRC (≥60 years), participants did not contributed person-time until age 60 years. We used the contrast test method to test whether the associations with risk factors differ between younger-onset and older-onset CRC.[37]

All the analyses were performed using SAS 9.4 (SAS Institute, Cary, NC, USA). Statistical tests used in the analysis all were 2-sided and a *P* value less than 0.05 was considered statistically significant.

Results

Among 45,351 men and 178,016 women in the study, followed for a median of 23 years, we documented 3058 CRC cases with available anatomic location data, including 474 cases in the cecum, 633 in the ascending colon, 250 in the transverse colon, 221 in the descending colon, 750 in the sigmoid colon, 202 in the rectosigmoid junction, and 528 in the rectum. Another 232 CRC cases without anatomic location information were censored at the time of diagnosis. Table 1 shows the basic characteristics of study participants. Tables 2–4 show the multivariable associations of risk factors with CRC across the seven anatomic subsites.

Demographic and clinical factors

The positive associations with CRC weakened from the cecum to rectum for age (HR per 5 years ranged from 1.62 to 1.32, $P_{\text{linear heterogeneity}}$=0.04) and family history of CRC (HR ranged from 1.86 to 1.10, $P_{\text{linear heterogeneity}}$=0.004) (Table 2). In contrast, the association strengthened for endoscopic screening from the cecum to rectum (HR ranged from 0.75 to 0.50, $P_{\text{linear heterogeneity}}$=0.005). A statistically significant overall heterogeneity was found for female sex ($P_{\text{overall heterogeneity}}$<0.001), with the highest HR observed for ascending colon cancer (1.73, 95% CI, 1.40–2.15) and the lowest HR for sigmoid colon cancer (0.54, 95% CI, 0.41–0.72).

Anthropometric and lifestyle factors

For anthropometric measures (Table 3), we did not find any statistically significant heterogeneity across the seven subsites, although in men, within distal colon cancer, a statistically significant heterogeneity was found for waist circumference between cancers in the descending and sigmoid colon (HR per 10 cm, 1.83 and 1.27, respectively, $P_{\text{heterogeneity}}$ =0.004). We found an increasing association for cancers from the cecum to transverse colon for alcohol intake (HR per 14 g/day, 0.99 to 1.25, $P_{\text{heterogeneity}}$= 0.02) and smoking before age 30 (HR per 20 pack-year, 1.08 to 1.43).

Dietary factors and dietary pattern

We found association with cancers strengthened from the cecum to rectum for whole grain (HR per 20 g/day ranged from 1.08 to 0.75, $P_{\text{linear heterogeneity}}$=0.02), cereal fiber (HR per 5g/day ranged from 1.13 to 0.60, $P_{\text{linear heterogeneity}}$=0.007), and processed red meat (HR per 3 serving/week ranged from 0.96 to 1.23, $P_{\text{linear heterogeneity}}$=0.04) (Table 4). A statistically significant heterogeneity was also found for folate intake between cancers in the descending and sigmoid colon (HR per 400pg/day, 1.10 and 0.76, respectively, $P_{\text{heterogeneity}}$=0.05). For EDIP, a particularly strong positive association was found with cancers in the cecum and sigmoid colon (HR per 1 unit, 1.54 and 1.52, respectively). In contrast, EDIH showed a particularly strong association with increased risk of transverse and descending colon cancers (HR per 1 unit, 2.19 and 2.02, respectively). A statistically significant heterogeneity was also found for EDIH between cancers in the rectosigmoid junction and rectum (HR per 1 unit, 0.71 and 1.54, respectively, $P_{\text{heterogeneity}}$=0.05).

Younger- and older-onset CRC

We identified 901 patients with younger-onset CRC and 2389 with older-onset CRC. The basic characteristics of patients in the two groups are summarized in Supplementary Table 1. As shown in Figure 1–2, we found that age and family history of CRC were more strongly associated with younger-onset CRC than older-onset CRC, whereas BMI, waist circumference and whole grain intake were more strongly associated with older-onset CRC. (All P for heterogeneity < 0.05) (Detailed data are presented in Supplementary Table 2–4). Height showed a stronger association with increased risk of younger- than older-onset CRC, whereas a stronger association with older- than younger-onset CRC was observed for several individual dietary factors, including cereal fiber, total vitamin D, total folate, and processed red meat. Similar patterns by tumor subsite were observed for younger- and older-onset CRC.

Discussion

Using data from three prospective cohorts, we evaluated 24 risk factors in relation to CRC risk by seven refined anatomic subsites. We found that the associations of certain risk factors with CRC varied substantially across refined subsites and even within the proximal, distal colon, and rectum. Our findings challenge the oversimplified classification of CRC into proximal and distal colon and rectal cancer and have implications for better understanding the etiology of CRC and improving prevention strategies.

Demographic factors

We found a weakened association of age with cancers from the cecum to rectum, with the strongest association observed in the ascending colon. Similarly, ascending colon cancer was more common in women, compared with other locations. These findings may be explained by the fact that older age and female sex are features of serrated CRC that is more commonly developed in the proximal colon, particularly ascending colon. Serrated CRC is characterized by certain molecular features, such as MSI, *BRAF* mutation, and CIMP.[39–42] Indeed, we found that ascending colon cancer had the highest rates of MSI, *BRAF* mutation, and CIMP. (Supplementary Figure 1).[13, 14] Interestingly, among subsites, we found that sigmoid colon cancer is particularly prevalent in men. This observation is consistent with the reported sex difference in incidence of CRC by anatomic sites.[43]

The positive association of family history with CRC risk was found to weaken gradually in cancers from the cecum to rectum. Consistent with our findings, Lynch syndrome in CRC has been associated with MSI-high status that is more prevalent in the proximal colon.[44, 45] Moreover, within microsatellite-stable CRC, family history was shown to be primarily associated with higher risk of CRC with low methylation level of long interspersed nuclear element-1 (LINE-1).[46] LINE-1 sequences constitute about 18% of the entire human genome and hypomethylation of LINE-1 has been suggested as a surrogate of genomic hypomethylation and been shown to vary across anatomic subsites of CRC.[14] Therefore, differences in heritable predisposition to epigenomic instability across the colorectum may be a potential explanation for the subsite heterogeneity in the familial risk of CRC.

Endoscopic screening and aspirin use

Endoscopic screening reduces CRC incidence and mortality.[47] In this study, we found that the beneficial association of endoscopic screening with CRC varied greatly by subsites. This is consistent with prior data indicating the limited effectiveness of endoscopic screening for prevention of proximal colon cancer. Possible explanations include anatomic and procedural (e.g., poor bowel preparation) reasons and overrepresentation in the proximal colon of serrated polyps that tend to be flat and pale and are difficult to detect and completely resect.[48]

We found that regular aspirin use was associated with lower risk of CRC across all anatomic locations, with no significant heterogeneity observed. Prior studies that assessed aspirin-CRC relationship by tumor subsites reported inconsistent findings. Some studies[49, 50] but not others[5,50–52] found a stronger association of aspirin use with proximal than distal colon cancer. On the other hand, the beneficial association of aspirin use with CRC risk has been primarily observed in tumors that overexpress PTGS2 (cyclooxygenase-2, COX-2),[51] whose expression does not seem to vary by CRC location.[52]

Lifestyle factors

We found that pack-years of smoking before age 30 and alcohol intake were more strongly associated with increased cancer risk in the transverse colon than other subsites. Prior studies have linked smoking and alcohol to specific molecular subtypes of CRC, namely *TP53, KRAS, BRAF* mutations, MSI, CIMP, and LINE-1.[30, 53–56] For example, alcohol consumption was associated with increased risk of CRC tumors characterized by LINE-1 hypomethylation cancers,[54] and some studies suggested smoking was associated with higher risks for CIMP-high, MSI-high and BRAF-mutated tumors.[30, 56] Therefore, differences in tumor molecular characteristics may underlie the subsite heterogeneity for the carcinogenic effect of smoking and alcohol.

Anthropometric factors

For anthropometric measures, consistent with prior data,[57] we found that the positive associations of BMI and waist circumference were stronger in men than in women, possibly due to the anti-CRC effect of elevated levels of estrogen that counteracts the adverse effect of adiposity in obese women.[58] Moreover, in line with prior studies that indicate a stronger association of BMI and waist circumference with distal than proximal colon cancer,[59] we found that these anthropometric measures were most strongly associated with increased risk of descending colon cancer in men. The biology underlying the observed differences across refined subsites remains poorly understood. Obesity-related hyperinsulinemia may promote carcinogenesis by increasing the bioavailability of tumor promoter IGF1 (insulin like growth factor 1). IGF1 has been shown to activate the PI3K/AKT signaling pathway by *PIK3CA* mutation.[60, 61] A much higher prevalence of *PIK3CA* mutation has been found in descending colon cancer than CRC in other locations.[11, 13] Moreover, increasing data indicate that obesity may influence cancer risk through alterations of the microbiome.[19] There are substantial differences in the microbial composition and function across subsites of the colorectum.[21, 62, 63] In addition, subsite differences in the biochemical environment,[64] mucosal immunology,[65] and gene expression may also play a role.[2, 66]

Dietary factors and dietary pattern

For whole grain and cereal fiber, we noted inverse association with cancers strengthened from the cecum to rectum. One prior study found that high intake of whole grains and dietary fiber were inversely associated with risk of ZP53-mutated CRC,[67] whose prevalence increases from the cecum to the rectum.[11] Moreover, compared to the colon, the rectum is much more exposed to genotoxic and cytotoxic damages due to the longer transit time and to the fecal mass storage before expulsion through defecation.[68] Increasing evidence indicates an intricate interplay between whole grain, fiber, gut microbiota, and CRC.[19, 69] For example, *Fusobacterium nucleatum* has been known to promote colorectal carcinogenesis through various mechanisms.[12, 70, 71] We recently reported higher intake of whole grains and cereal fiber was more strongly associated with lower risk of *F.nucleatum-positive* CRC than *F. nucleatum-negative* CRC.[72] However, it remains unknown how the gut microbiota in cancer-free individuals may modify the anti-CRC effect of fiber.

EDIP is an index that characterizes the inflammatory potential of diet based on circulating inflammatory markers. Higher EDIP has been associated with increased CRC risk.[73] In the current study, we found the positive associations between EDIP and risk of cecal and sigmoid cancer were 40–60% higher than cancers in other sites. The strong association of EDIP with cecal cancer may be related to functional interaction between *KRAS* mutation and inflammation. Several studies have reported a higher rate of *KRAS* mutation in cecal than non-cecal CRC.[11, 14, 74] There is evidence that mutant *KRAS* requires an additional, potentially inflammatory, stimulus to activate its oncogenic activity.[75, 76] On the other hand, gut microbiota may be another explanation for our findings due to the substantial differences in the composition of mucosa microbiota across subsites.[20]

Hyperinsulinemia and insulin resistance have been linked to increased risk of CRC.[36] In the current study, we found that EDIH, a dietary index that reflect insulinemic potential, was more strongly associated with increased risk of transverse and descending colon cancer. As discussed above, the IGF1-PI3K/AKT pathway may underlie the stronger associations of insulin-related factors with descending colon cancer. Regarding transverse colon cancer, its molecular characteristics are distinct from other right-sided locations and to a large degree resemble descending colon cancer.[11] However, because hyperinsulinemia and insulin resistance are related to complex metabolic changes, further studies are needed to better understand the molecular mechanisms underlying the strong association of EDIH with transverse and descending colon cancer.

Differences between younger- and older-onset CRC

The incidence of CRC in adults younger than 50 years has been increasing in the United States and several other countries. The increase is more pronounced for rectal cancer than colon cancer, and the causes for such increase remain unclear. In the current study, we found that family history of CRC were more strongly associated with increased risk of younger- and older-onset CRC, consistent with prior studies.[77] In contrast, adiposity and several dietary factors assessed in middle-to-late adulthood showed a stronger association with older- than younger-onset CRC. Interestingly we found that compared with older-onset CRC, attained height showed a stronger association with higher risk of younger-onset CRC,

especially rectal cancer. Height is influenced by genetic and early-life nutritional and health-related factors. Our previous study found that height as a marker of pre-adult IGF-I bioactivity was related to several Western-related cancers including CRC.[78] Therefore, these findings suggest a potential role of *in utero* and early-life nutritional exposures in young-onset CRC.

As the first effort to characterize the risk factor profiles of CRC according to refined anatomic subsites, our study has several strengths, including the prospective design, large sample size, repeated assessment of risk factors using validated instruments over 3 decades, and central medical record review for tumor location assessment. Our study also has some limitations. First, multiple comparisons were performed and thus some of the findings may be due to chance. Further studies may be needed for confirmation of the findings. However, all the risk factors and statistical comparisons were selected *a priori* on the basis of existing literature. We also interpret our results in a holistic way, prioritizing biological plausibility, coherence and consistency rather than statistical significance alone. Second, lifestyle and dietary factors are all self-reported and thus subject to measurement error. However, given the prospective design, any error in exposure assessment would have likely attenuated the observed associations. Third, despite the overall large size of the cohorts, the numbers of cases for certain subsites are rather limited. Finally, our participants are all health professionals and largely Caucasians, thus limiting the generalizability of our findings. However, our previously reported associations of risk factors with CRC have been largely confirmed by other cohorts. In addition, the homogenous study population helps reduce confounding, although residual confounding cannot be ruled out due to the observational design.

Our findings provide novel data for the heterogeneity of CRC risk factors according to tumor subsites. A better understanding of CRC risk factors according to tumor subsites may lead to development of subsite-specific prediction tools that may have better accuracy compared to the current prediction models for any CRCs. Accurate risk assessments may facilitate tailored screening strategies. Furthermore, identification of subsite-specific risk factors will facilitate development of targeted primary prevention strategies. For example, dietary modifications that improve metabolic health may be considered for individuals at high risk of developing distal colon cancer. While there is a long way ahead to realize these promises of precision prevention, our study provides the proof of principle for that future.

In conclusion, in this hypothesis-generating study, we found that the risk factor profiles differed for cancers across the colorectum, even within the proximal or distal tumors. Current proximal vs. distal classifications may not fully recapitulate the regional variations in the etiology of CRC. Future studies should account for precise anatomic subsites and elucidate the underlying mechanisms for the distinct risk factor profiles of cancer across the colorectum.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

We would like to thank the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data.

Funding: This work was supported by the American Cancer Society Mentored Research Scholar Grant (MRSG-17-220-01 - NEC to M.S.); by the U.S. National Institutes of Health (NIH) grants [P01 CA87969 to M.J. Stampfer; UM1 CA186107 to M.J. Stampfer; P01 CA55075, to W.C. Willett; U01 CA167552 to W.C. Willett; P50 CA127003 to C.S. Fuchs; K24 DK098311, R01 CA137178, R01 CA202704, R01 CA176726, to A.T.C.; R01 CA151993, R35 CA197735 to S.O.; R03 CA197879 to KW; R21 CA230873 to K.W. and S.O.; R00 CA215314 to M.S.]; and by grants from the American Institute for Cancer Research (K.W.), the Project P Fund for Colorectal Cancer Research, The Friends of the Dana-Farber Cancer Institute, Bennett Family Fund, and the Entertainment Industry Foundation through National Colorectal Cancer Research Alliance. Dr. Chan is a Stuart and Suzanne Steele MGH Research Scholar. The funders had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. L.W. was supported by the “3&3 Team Project” of the First Affiliated Hospital of Sun Yat-sen University.

Conflicts of interest: Andrew T. Chan previously served as a consultant for Bayer Pharma AG, Pfizer Inc., and Janssen Pharmaceuticals for work unrelated to the topic of this manuscript. This study was not funded by Bayer Pharma AG, Pfizer Inc., or Janssen Pharmaceuticals. No other conflict of interest exists.

Abbreviations

BMI	body mass index
CI	confidence interval
CIMP	CpG island methylator phenotype
EDIH	empirical dietary index for hyperinsulinemia
EDIP	empirical dietary inflammatory pattern
FFQ	food frequency questionnaire
HPFS	Health Professionals Follow-up Study
HR	hazard ratio
LINE-1	long interspersed nucleotide element-1
MET	metabolic equivalent of task
MSI	microsatellite instability
NHS	Nurses’ Health Study
PTGS2	(cyclooxygenase-2, COX-2)

Reference

1. Dienstmann R, Vermeulen L, Guinney J, et al. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. Nat Rev Cancer 2017;17:268.
2. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015;21:1350–6. [PubMed: 26457759]

3. Petrelli F, Tomasello G, Borgonovo K, et al. Prognostic Survival Associated With Left-Sided vs Right-Sided Colon Cancer: A Systematic Review and Meta-analysis. JAMA Oncol 2017;3:211–219. [PubMed: 27787550]
4. Buron Pust A, Alison R, Blanks R, et al. Heterogeneity of colorectal cancer risk by tumour characteristics: Large prospective study of UK women. Int J Cancer 2017;140:1082–1090. [PubMed: 27859268]
5. Murphy N, Ward HA, Jenab M, et al. Heterogeneity of Colorectal Cancer Risk Factors by Anatomical Subsite in 10 European Countries: A Multinational Cohort Study. Clin Gastroenterol Hepatol 2019;17:1323–1331 e6. [PubMed: 30056182]
6. Shin A, Joo J, Bak J, et al. Site-specific risk factors for colorectal cancer in a Korean population. PLoS One 2011;6:e23196. [PubMed: 21853085]
7. Demb J, Earles A, Martinez ME, et al. Risk factors for colorectal cancer significantly vary by anatomic site. BMJ Open Gastroenterol 2019;6:e000313.
8. Brandstedt J, Wangefjord S, Nodin B, et al. Gender, anthropometric factors and risk of colorectal cancer with particular reference to tumour location and TNM stage: a cohort study. Biol Sex Differ 2012;3:23. [PubMed: 23072404]
9. Benedix F, Schmidt U, Mroczkowski P, et al. Colon carcinoma--classification into right and left sided cancer or according to colonic subsite?--Analysis of 29,568 patients. Eur J Surg Oncol 2011;37:134–9. [PubMed: 21193285]
10. Jess P, Hansen IO, Gamborg M, et al. A nationwide Danish cohort study challenging the categorisation into right-sided and left-sided colon cancer. BMJ Open 2013;3.
11. Loree JM, Pereira AAL, Lam M, et al. Classifying Colorectal Cancer by Tumor Location Rather than Sidedness Highlights a Continuum in Mutation Profiles and Consensus Molecular Subtypes. Clin Cancer Res 2018;24:1062–1072. [PubMed: 29180604]
12. Mima K, Cao Y, Chan AT, et al. Fusobacterium nucleatum in Colorectal Carcinoma Tissue According to Tumor Location. Clin Transl Gastroenterol 2016;7:e200. [PubMed: 27811909]
13. Yamauchi M, Morikawa T, Kuchiba A, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. Gut 2012;61:847–54. [PubMed: 22427238]
14. Yamauchi M, Lochhead P, Morikawa T, et al. Colorectal cancer: a tale of two sides or a continuum? Gut 2012;61:794–7. [PubMed: 22490520]
15. Bae JM, Kim JH, Cho NY, et al. Prognostic implication of the CpG island methylator phenotype in colorectal cancers depends on tumour location. Br J Cancer 2013;109:1004–12. [PubMed: 23900220]
16. Ito M, Mitsuhashi K, Igarashi H, et al. MicroRNA-31 expression in relation to BRAF mutation, CpG island methylation and colorectal continuum in serrated lesions. Int J Cancer 2014;135:2507–15. [PubMed: 24752710]
17. Phipps AI, Chan AT, Ogino S. Anatomic subsite of primary colorectal cancer and subsequent risk and distribution of second cancers. Cancer 2013;119:3140–7. [PubMed: 23856984]
18. Purcell RV, Visnovska M, Biggs PJ, et al. Distinct gut microbiome patterns associate with consensus molecular subtypes of colorectal cancer. Sci Rep 2017;7:11590. [PubMed: 28912574]
19. Song M, Chan AT. Environmental Factors, Gut Microbiota, and Colorectal Cancer Prevention. Clin Gastroenterol Hepatol 2019;17:275–289. [PubMed: 30031175]
20. Zhang Z, Geng J, Tang X, et al. Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. ISME J 2014;8:881–93. [PubMed: 24132077]
21. Donaldson GP, Lee SM, Mazmanian SK. Gut biogeography of the bacterial microbiota. Nat Rev Microbiol 2016;14:20–32. [PubMed: 26499895]
22. Rimm EB, Giovannucci EL, Willett WC, et al. Prospective study of alcohol consumption and risk of coronary disease in men. Lancet 1991;338:464–8. [PubMed: 1678444]
23. Nimptsch K, Malik VS, Fung TT, et al. Dietary patterns during high school and risk of colorectal adenoma in a cohort of middle-aged women. Int J Cancer 2014;134:2458–67. [PubMed: 24493161]
24. Grodstein F, Martinez ME, Platz EA, et al. Postmenopausal hormone use and risk for colorectal cancer and adenoma. Ann Intern Med 1998;128:705–12. [PubMed: 9556463]

25. Ogino S, Kawasaki T, Kirkner GJ, et al. Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. J Mol Diagn 2007;9:305–14. [PubMed: 17591929]
26. Ogino S, Kawasaki T, Brahmandam M, et al. Sensitive sequencing method for KRAS mutation detection by Pyrosequencing. J Mol Diagn 2005;7:413–21. [PubMed: 16049314]
27. Ogino S, Kawasaki T, Brahmandam M, et al. Precision and performance characteristics of bisulfite conversion and real-time PCR (MethyLight) for quantitative DNA methylation analysis. J Mol Diagn 2006;8:209–17. [PubMed: 16645207]
28. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sports Exerc 1993;25:71–80. [PubMed: 8292105]
29. Chan AT, Giovannucci EL, Meyerhardt JA, et al. Long-term use of aspirin and nonsteroidal anti-inflammatory drugs and risk of colorectal cancer. JAMA 2005;294:914–23. [PubMed: 16118381]
30. Nishihara R, Morikawa T, Kuchiba A, et al. A prospective study of duration of smoking cessation and colorectal cancer risk by epigenetics-related tumor classification. Am J Epidemiol 2013;178:84–100. [PubMed: 23788674]
31. Yuan C, Spiegelman D, Rimm EB, et al. Validity of a Dietary Questionnaire Assessed by Comparison With Multiple Weighed Dietary Records or 24-Hour Recalls. Am J Epidemiol 2017;185:570–584. [PubMed: 28338828]
32. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;65:1220S–1228S; discussion 1229S-1231S. [PubMed: 9094926]
33. Tabung FK, Smith-Warner SA, Chavarro JE, et al. Development and Validation of an Empirical Dietary Inflammatory Index. J Nutr 2016;146:1560–70. [PubMed: 27358416]
34. Tabung FK, Smith-Warner SA, Chavarro JE, et al. An Empirical Dietary Inflammatory Pattern Score Enhances Prediction of Circulating Inflammatory Biomarkers in Adults. J Nutr 2017;147:1567–1577. [PubMed: 28659407]
35. Tabung FK, Wang W, Fung TT, et al. Development and validation of empirical indices to assess the insulinaemic potential of diet and lifestyle. Br J Nutr 2016:1–12.
36. Tabung FK, Wang W, Fung TT, et al. Association of dietary insulinemic potential and colorectal cancer risk in men and women. Am J Clin Nutr 2018;108:363–370. [PubMed: 29901698]
37. Wang M, Spiegelman D, Kuchiba A, et al. Statistical methods for studying disease subtype heterogeneity. Stat Med 2016;35:782–800. [PubMed: 26619806]
38. Wang M, Kuchiba A, Ogino S. A Meta-Regression Method for Studying Etiological Heterogeneity Across Disease Subtypes Classified by Multiple Biomarkers. Am J Epidemiol 2015;182:263–70. [PubMed: 26116215]
39. Chen D, Huang JF, Liu K, et al. BRAFV600E mutation and its association with clinicopathological features of colorectal cancer: a systematic review and meta-analysis. PLoS One 2014;9:e90607. [PubMed: 24594804]
40. Koo JH, Leong RW. Sex differences in epidemiological, clinical and pathological characteristics of colorectal cancer. J Gastroenterol Hepatol 2010;25:33–42. [PubMed: 19874446]
41. Young J, Simms LA, Biden KG, et al. Features of colorectal cancers with high-level microsatellite instability occurring in familial and sporadic settings: parallel pathways of tumorigenesis. Am J Pathol 2001;159:2107–16. [PubMed: 11733361]
42. Soreide K, Janssen EA, Soiland H, et al. Microsatellite instability in colorectal cancer. Br J Surg 2006;93:395–406. [PubMed: 16555243]
43. White A, Ironmonger L, Steele RJC, et al. A review of sex-related differences in colorectal cancer incidence, screening uptake, routes to diagnosis, cancer stage and survival in the UK. BMC Cancer 2018;18:906. [PubMed: 30236083]
44. Bapat B, Lindor NM, Baron J, et al. The association of tumor microsatellite instability phenotype with family history of colorectal cancer. Cancer Epidemiol Biomarkers Prev 2009;18:967–75. [PubMed: 19258475]
45. Chung DC, Rustgi AK. The hereditary nonpolyposis colorectal cancer syndrome: genetics and clinical implications. Ann Intern Med 2003;138:560–70. [PubMed: 12667026]

46. Ogino S, Nishihara R, Lochhead P, et al. Prospective study of family history and colorectal cancer risk by tumor LINE-1 methylation level. J Natl Cancer Inst 2013;105:130–40. [PubMed: 23175808]
47. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. N Engl J Med 2013;369:1095–105. [PubMed: 24047059]
48. Rosch T, Altenhofen L, Kretschmann J, et al. Risk of Malignancy in Adenomas Detected During Screening Colonoscopy. Clin Gastroenterol Hepatol 2018;16:1754–1761. [PubMed: 29902640]
49. Cook NR, Lee IM, Zhang SM, et al. Alternate-day, low-dose aspirin and cancer risk: long-term observational follow-up of a randomized trial. Ann Intern Med 2013;159:77–85. [PubMed: 23856681]
50. Rothwell PM, Wilson M, Elwin CE, et al. Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. Lancet 2010;376:1741–50. [PubMed: 20970847]
51. Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. N Engl J Med 2007;356:2131–42. [PubMed: 17522398]
52. Kawasaki T, Nosho K, Ohnishi M, et al. Cyclooxygenase-2 overexpression is common in serrated and non-serrated colorectal adenoma, but uncommon in hyperplastic polyp and sessile serrated polyp/adenoma. BMC Cancer 2008;8:33. [PubMed: 18230181]
53. Poynter JN, Haile RW, Siegmund KD, et al. Associations between smoking, alcohol consumption, and colorectal cancer, overall and by tumor microsatellite instability status. Cancer Epidemiol Biomarkers Prev 2009;18:2745–50. [PubMed: 19755657]
54. Schernhammer ES, Giovannucci E, Kawasaki T, et al. Dietary folate, alcohol and B vitamins in relation to LINE-1 hypomethylation in colon cancer. Gut 2010;59:794–9. [PubMed: 19828464]
55. Schernhammer ES, Ogino S, Fuchs CS. Folate and vitamin B6 intake and risk of colon cancer in relation to p53 expression. Gastroenterology 2008;135:770–80. [PubMed: 18619459]
56. Limsui D, Vierkant RA, Tillmans LS, et al. Cigarette smoking and colorectal cancer risk by molecularly defined subtypes. J Natl Cancer Inst 2010;102:1012–22. [PubMed: 20587792]
57. Abar L, Vieira AR, Aune D, et al. Height and body fatness and colorectal cancer risk: an update of the WCRF-AICR systematic review of published prospective studies. Eur J Nutr 2018;57:1701–1720. [PubMed: 29080978]
58. Giovannucci E Metabolic syndrome, hyperinsulinemia, and colon cancer: a review. Am J Clin Nutr 2007;86:s836–42. [PubMed: 18265477]
59. Laake I, Thune I, Selmer R, et al. A prospective study of body mass index, weight change, and risk of cancer in the proximal and distal colon. Cancer Epidemiol Biomarkers Prev 2010;19:1511–22. [PubMed: 20501754]
60. Huang XF, Chen JZ. Obesity, the PI3K/Akt signal pathway and colon cancer. Obes Rev 2009;10:610–6. [PubMed: 19527447]
61. Baba Y, Nosho K, Shima K, et al. Phosphorylated AKT expression is associated with PIK3CA mutation, low stage, and favorable outcome in 717 colorectal cancers. Cancer 2011;117:1399–408. [PubMed: 21425139]
62. Song M, Chan AT, Sun J. Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer. Gastroenterology 2020;158:322–340. [PubMed: 31586566]
63. Flynn KJ, Ruffin MTt, Turgeon DK, et al. Spatial Variation of the Native Colon Microbiota in Healthy Adults. Cancer Prev Res (Phila) 2018;11:393–402. [PubMed: 29636352]
64. Sandle GI. Salt and water absorption in the human colon: a modern appraisal. Gut 1998;43:294–9. [PubMed: 10189861]
65. Kirby JA, Bone M, Robertson H, et al. The number of intraepithelial T cells decreases from ascending colon to rectum. J Clin Pathol 2003;56:158.
66. Glebov OK, Rodriguez LM, Nakahara K, et al. Distinguishing right from left colon by the pattern of gene expression. Cancer Epidemiol Biomarkers Prev 2003;12:755–62. [PubMed: 12917207]
67. Slattery ML, Curtin K, Wolff RK, et al. Diet, physical activity, and body size associations with rectal tumor mutations and epigenetic changes. Cancer Causes Control 2010;21:1237–45. [PubMed: 20383576]

68. Gianfredi V, Nucci D, Salvatori T, et al. Rectal Cancer: 20% Risk Reduction Thanks to Dietary Fibre Intake. Systematic Review and Meta-Analysis. Nutrients 2019; 11.

69. Smith PM, Howitt MR, Panikov N, et al. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. Science 2013;341:569–73. [PubMed: 23828891]

70. Yang Y, Weng W, Peng J, et al. Fusobacterium nucleatum Increases Proliferation of Colorectal Cancer Cells and Tumor Development in Mice by Activating Toll-Like Receptor 4 Signaling to Nuclear Factor-kappaB, and Up-regulating Expression of MicroRNA-21. Gastroenterology 2017;152:851–866 e24. [PubMed: 27876571]

71. Yu T, Guo F, Yu Y, et al. Fusobacterium nucleatum Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. Cell 2017;170:548–563 e16. [PubMed: 28753429]

72. Mehta RS, Nishihara R, Cao Y, et al. Association of Dietary Patterns With Risk of Colorectal Cancer Subtypes Classified by Fusobacterium nucleatum in Tumor Tissue. JAMA Oncol 2017;3:921–927. [PubMed: 28125762]

73. Liu L, Nishihara R, Qian ZR, et al. Association Between Inflammatory Diet Pattern and Risk of Colorectal Carcinoma Subtypes Classified by Immune Responses to Tumor. Gastroenterology 2017;153:1517–1530 e14. [PubMed: 28865736]

74. Landau MA, Zhu B, Akwuole FN, et al. Site-specific Differences in Colonic Adenocarcinoma: KRAS Mutations and High Tumor Budding Are More Frequent in Cecal Adenocarcinoma. Am J Surg Pathol 2018;42:351–358. [PubMed: 29240583]

75. Philip B, Roland CL, Daniluk J, et al. A high-fat diet activates oncogenic Kras and COX2 to induce development of pancreatic ductal adenocarcinoma in mice. Gastroenterology 2013;145:1449–58. [PubMed: 23958541]

76. Huang H, Daniluk J, Liu Y, et al. Oncogenic K-Ras requires activation for enhanced activity. Oncogene 2014;33:532–5. [PubMed: 23334325]

77. Gausman V, Dornblaser D, Anand S, et al. Risk Factors Associated With Early-Onset Colorectal Cancer. Clin Gastroenterol Hepatol 2019.

78. Giovannucci E, Rimm EB, Liu Y, et al. Height, predictors of C-peptide and cancer risk in men. Int J Epidemiol 2004;33:217–25. Author names in bold designate shared co-first authorship [PubMed: 15075172]

What you need to know

Background and Context

The molecular features of colorectal tumors differ with their anatomic location, usually classified as proximal or distal. It is important to identify risk factors for colorectal cancer (CRC) at refined anatomic subsites.

New Findings

Risk factors, including age, family history, diet, screening, alcohol use, smoking, and sex, differ for cancers along the colorectum.

Limitations

This was a study of 3 large cohorts in the United States—additional studies are needed of other populations.

Impact

The proximal vs distal colon classifications are too broad for accurate calculation of CRC risk; risk factors vary for tumors in different regions of the colorectum.

Lay Summary

Risk factors differ for tumors that develop in different regions of the colon and rectum.

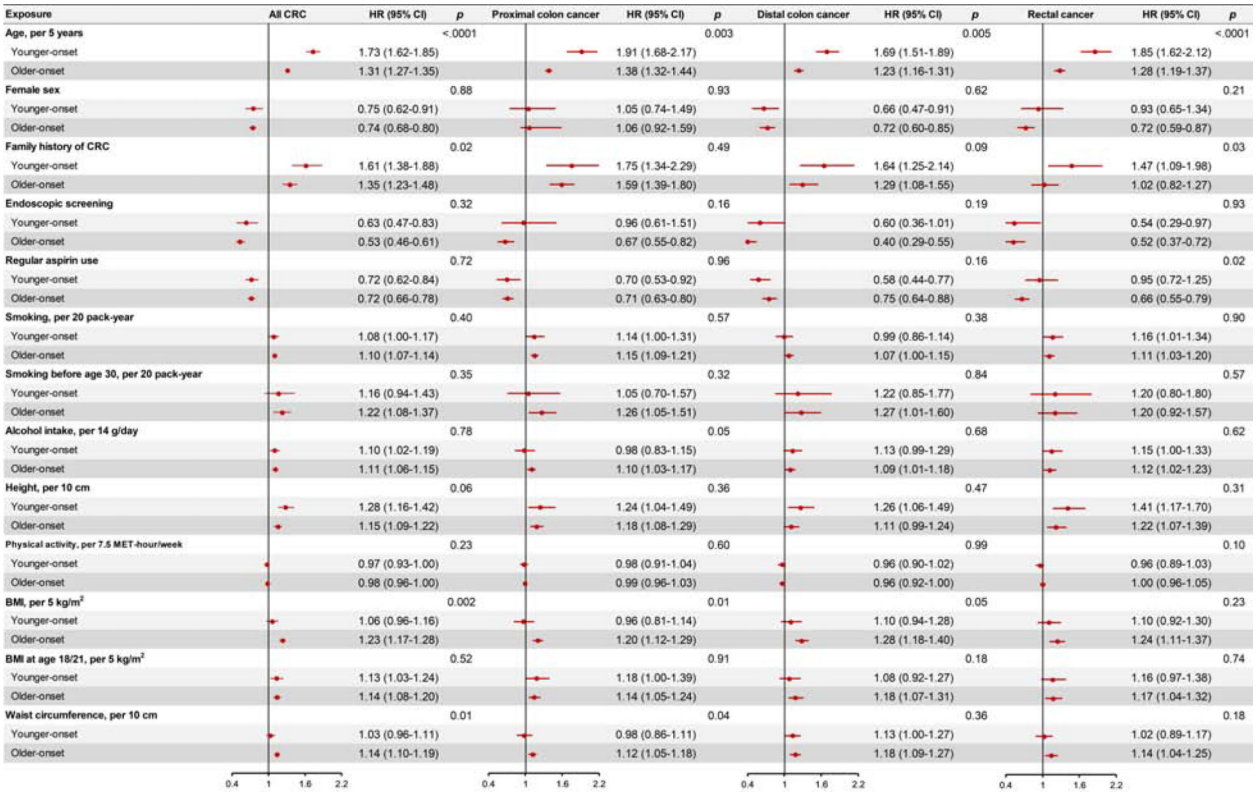


Figure 1.

Multivariable associations of demographic, clinical and lifestyle factors with subsite-specific risk of colorectal cancer according to age at diagnosis (Younger-onset CRC: diagnosed at age of <60 years; older-onset CRC: diagnosed at age of ≥60 years) in the NHS, NHS2, and HPFS.

Hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated using age- and cohort-stratified Cox proportional hazards model with further adjustment for race, height (continuous), family history of colorectal cancer (yes or no), history of lower gastrointestinal endoscopic screening (yes or no), body mass index (continuous), pack-years of smoking (continuous), physical activity (continuous), alcohol intake (continuous), and regular aspirin use (yes or no). When age and sex are the main exposures, the model was only adjusted for each other of the two variables. *P* for heterogeneity was calculated between younger-onset CRC and older-onset CRC using the contrast test method. Abbreviations: BMI, body mass index; CRC, colorectal cancer; HPFS, the Health Professionals Follow-up Study; NHS, the Nurses' Health Study.

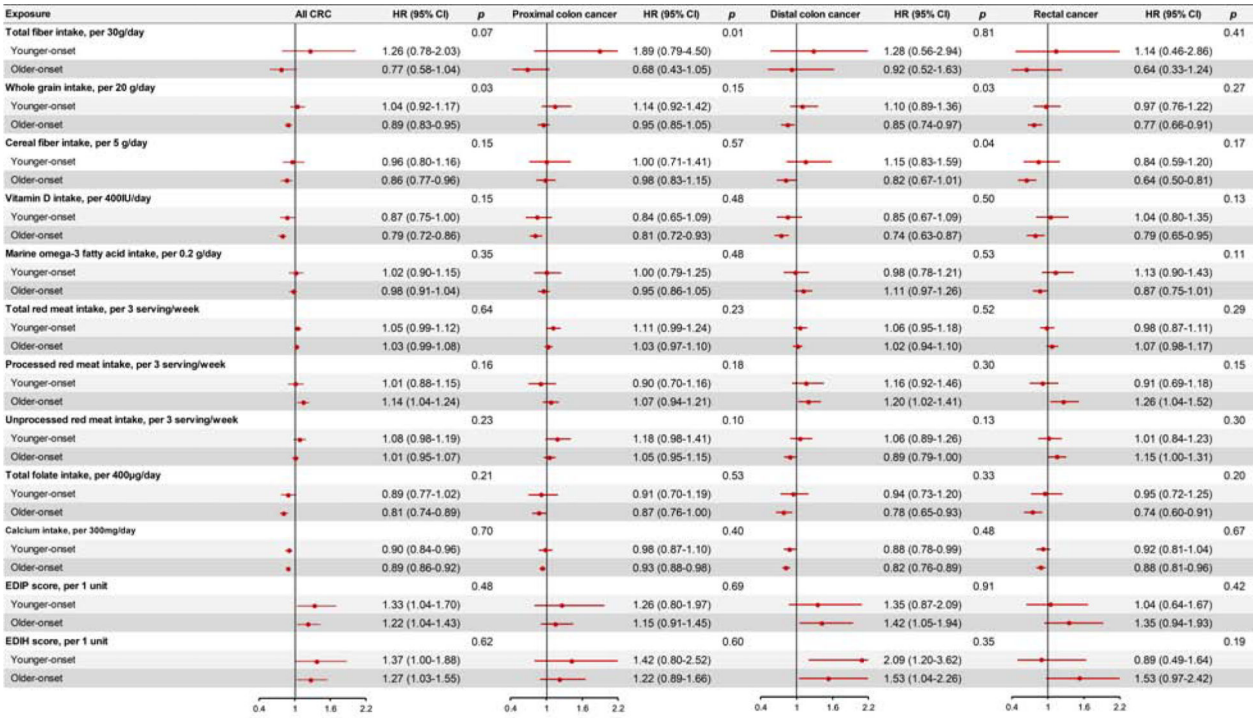


Figure 2.

Multivariable associations of dietary factors with subsite-specific risk of colorectal cancer according to age at diagnosis (Younger-onset CRC: diagnosed at age of <60 years; older-onset CRC: diagnosed at age of ≥60 years) in the NHS, NHS2, and HPFS.

Hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated using age- and cohort-stratified Cox proportional hazards model with further adjustment for race, height (continuous), family history of colorectal cancer (yes or no), history of lower gastrointestinal endoscopic screening (yes or no), body mass index (continuous), pack-years of smoking (continuous), physical activity (continuous), alcohol intake (continuous), and regular aspirin use (yes or no). *P* for heterogeneity was calculated between younger-onset CRC and older-onset CRC using the contrast test method.

Abbreviations: CRC, colorectal cancer; EDIP, empirical dietary inflammatory pattern; EDIH, empirical dietary index for hyperinsulinemia; HPFS, the Health Professionals Follow-up Study; NHS, the Nurses' Health Study.

Table 1

Basic characteristics of study participants in the NHS, NHS2, and HPFS cohorts [a].

Risk factors	Participants without CRC	Participants with CRC						
		Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectosigmoid	Rectum
	n=220309	n=474	n=633	n=250	n=221	n=750	n=202	n=528
Age [b], years	56.6 (12.8)	60.0 (10.4)	61.1 (10.4)	60.7 (10.1)	59.0 (10.9)	58.4 (10.7)	57.5 (10.7)	58.4 (10.9)
Whites, %	97	98	96	97	98	97	97	99
Female, %	82	66	82	61	74	68	74	69
Family history of CRC, %	13	15	23	25	11	17	16	15
Regular aspirin use, % [c]	30	29	25	23	25	28	28	27
Height, cm	167 (8)	167 (8)	167 (8)	169 (8)	168 (7)	167 (8)	169 (8)	168 (9)
Body mass index, kg/m^2	26.2 (5.4)	25.5 (4.7)	26.8 (5.7)	26.4 (4.7)	26.8 (5.4)	26.6 (5.1)	25.8 (5.4)	26.4 (5.6)
BMI at age 18/21, kg/m^2	21.5 (3.4)	21.6 (3.4)	22.0 (3.6)	21.7 (2.8)	21.7 (3.6)	21.8 (3.8)	21.4 (3.6)	21.8 (3.4)
Waist circumference, cm	84.8 (13.7)	84.0 (12.5)	86.9 (14.4)	85.8 (12.3)	88.3 (15.1)	86.0 (12.1)	86.9 (13.9)	85.3 (12.3)
Pack-years of smoking	9.9 (16.6)	11.9 (18.9)	11.6 (17.8)	12.5 (18.8)	13.2 (18.1)	11.2 (17.8)	12.1 (17.7)	12.5 (18.3)
Pack-years of smoking before age 30	3.7 (5.7)	3.9 (5.4)	3.8 (5.5)	4.3 (6.5)	5.1 (6.2)	3.9 (5.9)	4.2 (5.7)	4.0 (5.3)
Physical activity, MET-hours/week [d]	21.2 (21.1)	17.5 (15.9)	20.2 (22.9)	18.1 (13.2)	19.0 (20.6)	19.9 (17.9)	19.4 (20.0)	20.9 (26.5)
Alcohol intake, g/day	6.1 (9.7)	6.0 (9.7)	6.0 (9.8)	8.3 (13.2)	6.6 (11.3)	6.3 (10.2)	6.5 (10.8)	7.2 (11.5)
Total folate intake, μg/day	488 (237)	440 (218)	466 (230)	431 (305)	467 (243)	428 (208)	453 (214)	445 (234)
Calcium intake, mg/day	1006 (395)	920 (346)	990 (418)	891 (375)	876 (320)	908 (367)	959 (375)	937 (376)
Vitamin D intake, IU/day	395 (237)	353 (211)	364 (213)	348 (292)	349 (279)	342 (215)	365 (225)	362 (236)
Marine omega-3 fatty acid intake, g/day	0.21 (0.18)	0.19 (0.18)	0.20 (0.16)	0.22 (0.17)	0.21 (0.15)	0.21 (0.18)	0.18 (0.16)	0.20 (0.20)
Total red meat intake, serving/week	6.2 (3.7)	6.2 (3.6)	6.3 (3.8)	7.2 (3.8)	7.0 (4.2)	6.5 (4.3)	6.1 (3.6)	6.5 (3.9)
Processed red meat intake, serving/week	1.9 (1.9)	1.9 (1.9)	2.0 (1.9)	2.3 (2.3)	2.0 (2.3)	2.1 (2.4)	1.9 (1.9)	2.0 (2.1)
Unprocessed red meat intake, serving/week	3.8 (2.4)	3.9 (2.3)	3.9 (2.5)	4.2 (2.4)	4.3 (2.8)	3.9 (2.6)	3.7 (2.3)	4.0 (2.5)
Total fiber intake, g/day	18.4 (5.7)	18.1 (5.8)	18.7 (5.8)	17.1 (5.1)	18.5 (5.7)	18.6 (6.0)	18.8 (6.1)	18.1 (5.8)
Whole grain intake, g/day	21.7 (14.7)	20.4 (14.8)	22.2 (16.1)	18.8 (12.0)	21.2 (14.8)	20.0 (14.8)	23.4 (18.7)	19.4 (15.4)
Cereal fiber intake, g/day	5.24 (2.72)	5.01 (2.59)	5.33 (3.06)	4.62 (2.05)	5.05 (2.53)	4.94 (2.66)	5.51 (3.18)	4.74 (2.62)
Empirical dietary inflammatory pattern score	0 (0.34)	0.02 (0.31)	0.02 (0.33)	0.03 (0.33)	0.06 (0.38)	0.04 (0.35)	−0.01 (0.30)	−0.01 (0.36)

Risk factors	Participants without CRC	Participants with CRC						
		Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectosigmoid	Rectum
	n=220309	n=474	n=633	n=250	n=221	n=750	n=202	n=528
Empirical dietary index for hyperinsulinemia	0.38 (0.25)	0.38 (0.25)	0.39 (0.27)	0.42 (0.24)	0.45 (0.31)	0.40 (0.25)	0.38 (0.25)	0.39 (0.27)

Abbreviations: CRC, colorectal cancer; HPFS, the Health Professionals Follow-up Study; MET, metabolic equivalent task; NHS, the Nurses' Health Study.

[a]Data are based on repeatedly collected information for each participant up to CRC diagnosis (for cases) or the end of follow-up (for non-cases). All variables are adjusted for age and sex except for age and sex themselves. Cumulative average values across follow-up are presented. Mean (standard deviation) is presented for continuous variables and percentage for categorical variables.

[b]Mean age across follow-up.

[c]A standard tablet contains 325 mg aspirin, and regular users were defined as those who used at least two standard tablets per week.

[d]Physical activity is calculated by the product sum of the MET of each specific recreational activity and hours spent on that activity per week. For physical activity, the follow-up started in 1986 in NHS.

Table 2

Multivariable associations of demographic and clinical factors with risk of colorectal cancer according to refined subsites in the NHS, NHS2, and HPFS[a].

Risk factors	Proximal colon cancer				Distal colon cancer			Rectum cancer			*P for overall heterogeneity*[b]	*P for linear heterogeneity*[c]
	Cecum	Ascending colon	Transverse colon	Overall	Descending colon	Sigmoid colon	Overall	Rectosigmoid	Rectum	Overall		
	n=474	n=633	n=250	n=1357	n=221	n=750	n=971	n=202	n=528	n=730		
Age, per 5 years												
HR	1.43	1.62	1.45	1.51	1.42	1.39	1.39	1.32	1.37	1.36		
95% CI	1.35–1.52	1.53–1.71	1.34–1.57	1.45–1.56	1.3–1.55	1.32–1.45	1.34–1.45	1.2–1.44	1.29–1.45	1.29–1.42		
P_{trend}	1.43	1.62	1.45	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.04
P for heterogeneity within major subsites[b]				0.006			0.72			0.52		
Female sex												
HR	0.63	1.73	0.61	0.93	0.83	0.54	0.65	0.84	0.68	0.70		
95% CI	0.51–0.78	1.40–2.15	0.51–0.72	0.81–1.05	0.6–1.16	0.41–0.72	0.56–0.76	0.61–1.17	0.56–0.84	0.59–0.84		
P	<.0001	<.0001	<.0001	0.24	0.27	<.0001	<.0001	0.31	0.0003	<.0001	<0.001	0.70
P for heterogeneity within main subsites[b]				<0.001			0.10			0.29		
Family history of colorectal cancer												
HR	1.55	1.66	1.86	1.65	1.22	1.46	1.41	1.26	1.10	1.14		
95%CI	1.25–1.91	1.39–1.98	1.41–2.46	1.46–1.87	0.88–1.70	1.23–1.73	1.21–1.64	0.89–1.79	0.88–1.37	0.94–1.37		
P	<0.001	<0.001	<0.001	<0.001	0.24	<.0001	<0.001	0.19	0.42	0.17	0.02	0.004
P for heterogeneity within main subsites[b]				0.57			0.37			0.52		
Endoscopic screening												
HR	0.75	0.71	0.67	0.72	0.45	0.36	0.38	0.53	0.50	0.51		
95% CI	0.54–1.03	0.53–0.94	0.41–1.08	0.59–0.87	0.24–0.83	0.25–0.51	0.28–0.51	0.30–0.94	0.34–0.72	0.37–0.69		
P	0.07	0.02	0.10	<0.001	0.01	<0.001	<0.001	0.03	<0.001	<0.001	0.02	0.005
P for heterogeneity within main subsites[b]				0.94			0.62			0.83		
Regular aspirin use												

Risk factors	Proximal colon cancer				Distal colon cancer			Rectum cancer			*P for overall heterogeneity*[b]	*P for linear heterogeneity*[c]
	Cecum	Ascending colon	Transverse colon	Overall	Descending colon	Sigmoid colon	Overall	Rectosigmoid	Rectum	Overall		
	n=474	n=633	n=250	n=1357	n=221	n=750	n=971	n=202	n=528	n=730		
HR	0.76	0.69	0.67	0.71	0.62	0.73	0.70	0.65	0.81	0.76		
95% CI	0.63–0.92	0.58–0.81	0.51–0.87	0.63–0.79	0.46–0.83	0.62–0.85	0.61–0.80	0.47–0.88	0.68–0.98	0.65–0.89		
P	0.005	<0.001	0.003	<0.001	<0.001	<0.001	<0.001	0.005	0.03	<0.001	0.65	0.51
P for heterogeneity within main subsites[b]				0.76			0.39			0.14		

[a]Age- and cohort-stratified Cox proportional hazards model was used with further adjustment for race, height (continuous), family history of colorectal cancer (yes or no), history of lower gastrointestinal endoscopic screening (yes or no), body mass index (continuous), pack-years of smoking (continuous), physical activity (continuous), alcohol intake (continuous), and regular aspirin use (yes or no). When age and sex are the main exposures, the model was only adjusted for each other of the two variables.

[b]Overall heterogeneity was tested by using the meta-regression method with a subsite-specific random effect term as a nominal variable.

[c]Linear heterogeneity was tested by using the meta-regression method with a subsite-specific random effect term as an ordinal variable.

Table 3

Multivariable associations of lifestyle and anthropometric factors with risk of colorectal cancer according to refined subsites in the NHS, NHS2, and HPFS [a].

Risk factor	Proximal colon cancer				Distal colon cancer			Rectum cancer			*P for overall heterogeneity* [c]	*P for linear heterogeneity* [d]
	Cecum	Ascending colon	Transverse colon	Overall	Descending colon	Sigmoid colon	Overall	Rectosigmoid	Rectum	Overall		
	n=474	n=633	n=250	n=1357	n=221	n=750	n=971	n=202	n=528	n=730		
Smoking												
Packyears of smoking, per 20 pack-year												
HR	1.12	1.14	1.12	1.13	1.10	1.03	1.05	1.11	1.12	1.11		
95% CI	1.03–1.22	1.06–1.22	1.00–1.26	1.07–1.19	0.97–1.25	0.96–1.11	0.98–1.12	0.96–1.27	1.03–1.21	1.04–1.20		
P_{trend}	0.01	0.0005	0.05	<0.001	0.13	0.43	0.15	0.15	0.009	0.003	0.61	0.40
P for heterogeneity within main subsites [c]				0.80			0.25			0.77		
Packyears of smoking before age 30, per 20 pack-year												
HR	1.08	1.27	1.43	1.21	1.37	1.19	1.23	1.27	1.26	1.17		
95% CI	0.81–1.46	0.98–1.65	0.99–2.07	1.01–1.43	1.06–2.33	0.95–1.49	1.00–1.50	0.81–1.97	0.96–1.64	0.93–1.48		
P_{trend}	0.60	0.07	0.06	0.03	0.03	0.13	0.048	0.30	0.10	0.18	0.67	0.62
P for heterogeneity within main subsites [c]				0.10			0.21			0.76		
Alcohol intake, per 14 g/day												
HR	0.99	1.02	1.25	1.06	1.18	1.11	1.12	1.10	1.14	1.12		
95% CI	0.89–1.11	0.92–1.13	1.11–1.41	0.99–1.13	1.02–1.37	1.02–1.20	1.05–1.20	0.93–1.30	1.04–1.26	1.03–1.22		
P_{trend}	0.91	0.71	<0.001	0.10	0.02	0.01	0.001	0.27	0.008	0.01	0.11	0.09
P for heterogeneity within main subsites [c]				0.02			0.32			0.66		
BMI, per 5 kg/m^2												
Men	n=197	n=143	n=107	n=447	n=70	n=292	n=362	n=68	n=186	n=254		
HR	1.21	1.32	1.16	1.24	1.49	1.28	1.32	1.14	1.33	1.25		
95% CI	0.97–1.50	1.03–1.68	0.87–1.55	1.08–1.43	1.13–1.96	1.11–1.49	1.16–1.50	0.80–1.64	1.08–1.63	1.05–1.49		

Risk factor	Proximal colon cancer				Distal colon cancer			Rectum cancer			*P for overall heterogeneity c*	*P for linear heterogeneity d*
	Cecum	Ascending colon	Transverse colon	Overall	Descending colon	Sigmoid colon	Overall	Rectosigmoid	Rectum	Overall		
	n=474	n=633	n=250	n=1357	n=221	n=750	n=971	n=202	n=528	n=730		
P_{trend}	0.09	0.03	0.31	0.003	0.005	0.001	<0.001	0.47	0.008	0.01	0.69	0.70
P for heterogeneity within main subsites [c]				0.87			0.21			0.62		
Women	n=277	n=490	n=143	n=910	n=151	n=458	n=609	n=134	n=342	n=476		
HR	1.11	1.13	1.27	1.15	1.12	1.21	1.19	1.06	1.24	1.19		
95% CI	0.97–1.28	1.01–1.25	1.06–1.52	1.06–1.24	0.92–1.37	1.09–1.35	1.09–1.31	0.85–1.32	1.09–1.40	1.07–1.32		
P_{trend}	0.14	0.03	0.01	<0.001	0.25	0.0003	<0.001	0.59	<0.001	0.002	0.63	0.29
P for heterogeneity within main subsites [c]				0.68			0.33			0.22		
BMI at age 18/21, per 5 kg/m2												
Men	n=197	n=143	n=107	n=447	n=70	n=292	n=362	n=68	n=186	n=254		
HR	1.11	1.02	1.08	1.07	0.94	1.12	1.08	1.00	1.12	1.06		
95% CI	0.93–1.32	0.85–1.22	0.86–1.35	0.96–1.20	0.75–1.18	0.98–1.29	0.96–1.21	0.77–1.30	0.94–1.33	0.92–1.22		
P_{trend}	0.23	0.85	0.50	0.21	0.60	0.10	0.22	0.99	0.22	0.43	0.85	0.81
P for heterogeneity within main subsites [c]				0.84			0.20			0.43		
Women	n=277	n=490	n=143	n=910	n=151	n=458	n=609	n=134	n=342	n=476		
HR	1.09	1.24	1.33	1.20	1.22	1.21	1.21	1.26	1.25	1.25		
95% CI	0.90–1.32	1.07–1.42	1.03–1.71	1.09–1.33	0.96–1.55	1.05–1.40	1.07–1.37	0.98–1.62	1.07–1.47	1.09–1.44		
P_{trend}	0.39	0.003	0.03	<0.001	0.11	0.007	0.002	0.07	0.006	0.001	0.98	0.50
P for heterogeneity within main subsites [c]				0.70			0.94			0.96		
Waist circumference, per 10 cm												
Men	n=197	n=143	n=107	n=447	n=70	n=292	n=362	n=68	n=186	n=254		
HR	1.17	1.31	1.09	1.21	1.83	1.27	1.37	1.31	1.25	1.27		
95% CI	0.99–1.39	1.09–1.57	0.86–1.38	1.08–1.34	1.43–2.33	1.12–1.44	1.22–1.53	1.02–1.69	1.04–1.49	1.10–1.47		
P_{trend}	0.06	0.004	0.47	<0.001	<0.001	<0.001	<0.001	0.04	0.02	0.001	0.09	0.58
P for heterogeneity within main subsites [c]				0.44			0.004			0.89		

Risk factor	Proximal colon cancer				Distal colon cancer			Rectum cancer			*P for overall heterogeneity c*	*P for linear heterogeneity d*
	Cecum	Ascending colon	Transverse colon	Overall	Descending colon	Sigmoid colon	Overall	Rectosigmoid	Rectum	Overall		
	n=474	n=633	n=250	n=1357	n=221	n=750	n=971	n=202	n=528	n=730		
Women	n=277	n=490	n=143	n=910	n=151	n=458	n=609	n=134	n=342	n=476		
HR	0.98	1.08	1.14	1.06	1.02	1.09	1.08	1.04	1.07	1.06		
95% CI	0.87–1.11	0.99–1.18	0.97–1.34	0.99–1.13	0.87–1.21	0.99–1.20	0.99–1.17	0.88–1.23	0.96–1.19	0.97–1.16		
P_{trend}	0.72	0.08	0.10	0.08	0.79	0.07	0.08	0.62	0.24	0.22	0.70	0.59
P for heterogeneity within main subsites [c]				0.26			0.54			0.69		
Physical activity, per 7.5 MET-hours/week [b]												
HR	0.99	1.01	0.97	0.99	0.96	0.96	0.96	0.97	0.98	0.98		
95% CI	0.94–1.04	0.97–1.06	0.90–1.04	0.96–1.02	0.89–1.03	0.93–1.00	0.93–1.00	0.90–1.05	0.93–1.03	0.94–1.02		
P_{trend}	0.58	0.63	0.35	0.65	0.25	0.07	0.04	0.43	0.42	0.31	0.49	0.35
P for heterogeneity within main subsites [c]				0.96			0.76			0.43		
Height, per 10 cm												
HR	1.14	1.25	1.23	1.20	1.20	1.10	1.12	1.21	1.27	1.25		
95% CI	0.99–1.31	1.10–1.42	1.01–1.49	1.10–1.30	0.97–1.47	0.99–1.23	1.02–1.24	0.97–1.50	1.11–1.46	1.12–1.40		
P_{trend}	0.08	<0.001	0.04	<0.001	0.09	0.08	0.02	0.09	<0.001	<0.001	0.60	0.77
P for heterogeneity within main subsites [c]				0.56			0.48			0.66		

[a] Age- and cohort-stratified Cox proportional hazards model was used with adjustment for age (continuous), race, height (continuous), family history of colorectal cancer (yes or no), history of lower gastrointestinal endoscopic screening (yes or no), body mass index (continuous), pack-years of smoking (continuous), physical activity (continuous), alcohol intake (continuous), and regular aspirin use (yes or no).

[b] For physical activity, the follow-up started in 1986 in NHS.

[c] Heterogeneity was tested by using meta-regression method with a subsite-specific random effect term as a nominal variable.

[d] Linear heterogeneity was tested by using meta-regression method with a subsite-specific random effect term as an ordinal variable.

Table 4

Multivariable associations of dietary factors with risk of colorectal cancer according to refined subsites in the NHS, NHS2, and HPFS.

Risk factor	Proximal colon cancer				Distal colon cancer			Rectum cancer			*P for overall heterogeneity c*	*P for linear heterogeneity d*
	Cecum	Ascending colon	Transverse colon	Overall	Descending colon	Sigmoid colon	Overall	Rectosigmoid	Rectum	Overall		
	n=474	n=633	n=250	n=1357	n=221	n=750	n=971	n=202	n=528	n=730		
Total fiber intake, per 30g/day												
HR	0.93	1.09	0.60	0.92	1.36	0.86	0.95	1.55	0.59	0.77		
95% CI	0.46–1.86	0.60–2.01	0.23–1.58	0.61–1.39	0.49–3.72	0.49–1.48	0.58–1.54	0.55–4.42	0.31–1.14	0.45–1.34		
P_{trend}	0.83	0.77	0.30	0.70	0.55	0.58	0.83	0.41	0.12	0.36	0.30	0.45
P for heterogeneity within main subsites [c]				0.18			0.76			0.12		
Whole grain intake, per 20 g/day												
HR	1.08	0.95	0.87	0.98	1.00	0.91	0.93	1.07	0.75	0.82		
95% CI	0.91–1.27	0.82–1.11	0.69–1.10	0.88–1.08	0.78–1.28	0.79–1.04	0.82–1.04	0.82–1.38	0.63–0.88	0.71–0.94		
P_{trend}	0.39	0.54	0.23	0.65	0.99	0.16	0.20	0.63	<0.001	0.005	0.03	0.02
P for heterogeneity within main subsites [c]				0.17			0.80			0.03		
Cereal fiber intake, per 5 g/day												
HR	1.13	0.96	0.84	0.99	0.95	0.92	0.93	1.11	0.60	0.71		
95% CI	0.87–1.46	0.76–1.20	0.58–1.21	0.85–1.16	0.64–1.40	0.75–1.13	0.77–1.11	0.75–1.65	0.46–0.77	0.57–0.88		
P_{trend}	0.36	0.70	0.34	0.92	0.79	0.45	0.42	0.59	<0.001	0.001	0.003	0.007
P for heterogeneity within main subsites [c]				0.18			0.83			0.01		
Vitamin D intake, per 400IU/day												
HR	0.87	0.81	0.70	0.80	0.81	0.76	0.77	0.89	0.84	0.87		
95% CI	0.71–1.06	0.67–0.96	0.53–0.93	0.71–0.91	0.60–1.09	0.64–0.89	0.66–0.88	0.66–1.21	0.69–1.02	0.74–1.03	0.71	0.97
P_{trend}	0.16	0.02	0.02	<0.001	0.16	<0.001	<0.001	0.47	0.07	0.10		
P for heterogeneity within main subsites [c]				0.30			0.84			0.78		
Marine omega-3 fatty acid intake, per 0.2 g/day												

Risk factor	Proximal colon cancer				Distal colon cancer			Rectum cancer			*P for overall heterogeneity c*	*P for linear heterogeneity d*
	Cecum	Ascending colon	Transverse colon	Overall	Descending colon	Sigmoid colon	Overall	Rectosigmoid	Rectum	Overall		
	n=474	n=633	n=250	n=1357	n=221	n=750	n=971	n=202	n=528	n=730		
HR	0.90	0.97	1.02	0.95	1.09	1.08	1.08	0.81	0.99	0.94		
95% CI	0.76–1.05	0.84–1.12	0.82–1.27	0.86–1.05	0.86–1.38	0.95–1.23	0.97–1.21	0.63–1.04	0.85–1.15	0.82–1.07		
P_{trend}	0.18	0.67	0.85	0.30	0.47	0.22	0.16	0.09	0.86	0.33	0.37	0.49
P for heterogeneity within main subsites [c]				0.59			0.94			0.19		
Total red meat intake, per 3 serving/week												
HR	1.02	1.04	1.14	1.05	1.05	1.04	1.04	0.97	1.07	1.05		
95% CI	0.93–1.12	0.96–1.13	1.00–1.29	1.00–1.11	0.92–1.20	0.96–1.11	0.97–1.11	0.84–1.12	0.98–1.17	0.97–1.13		
P_{trend}	0.70	0.36	0.046	0.07	0.47	0.34	0.24	0.65	0.11	0.22		
P for heterogeneity within main subsites [c]				0.19			0.72			0.15	0.46	0.82
Processed red meat intake, per 3 serving/week												
HR	0.96	0.97	1.20	1.01	1.23	1.17	1.19	1.01	1.21	1.16		
95% CI	0.78–1.17	0.81–1.16	0.92–1.57	0.90–1.14	0.92–1.64	1.00–1.37	1.03–1.36	0.73–1.37	1.00–1.46	0.99–1.36		
P_{trend}	0.68	0.72	0.18	0.85	0.16	0.04	0.01	0.98	0.04	0.07	0.12	0.04
P for heterogeneity within main subsites [c]				0.18			0.64			0.23		
Unprocessed red meat intake, per 3 serving/week												
HR	1.06	1.07	1.16	1.08	0.95	0.95	0.95	1.01	1.14	1.11		
95% CI	0.92–1.22	0.94–1.21	0.96–1.41	1.00–1.18	0.77–1.18	0.85–1.06	0.86–1.05	0.81–1.26	1.00–1.30	0.99–1.24		
P_{trend}	0.46	0.32	0.13	0.06	0.66	0.35	0.30	0.91	0.05	0.08	0.31	0.96
P for heterogeneity within main subsites [c]				0.19			0.87			0.27		
Total folate intake, per 400μg/day												
HR	0.94	0.89	0.68	0.86	1.10	0.76	0.83	0.73	0.85	0.82		
95% CI	0.76–1.16	0.73–1.07	0.50–0.92	0.76–0.98	0.81–1.49	0.64–0.91	0.71–0.96	0.52–1.01	0.69–1.04	0.69–0.98		
P_{trend}	0.57	0.20	0.01	0.02	0.56	0.002	0.01	0.06	0.12	0.03	0.17	0.30

Risk factor	Proximal colon cancer				Distal colon cancer			Rectum cancer			*P for overall heterogeneity c*	*P for linear heterogeneity d*
	Cecum	Ascending colon	Transverse colon	Overall	Descending colon	Sigmoid colon	Overall	Rectosigmoid	Rectum	Overall		
	n=474	n=633	n=250	n=1357	n=221	n=750	n=971	n=202	n=528	n=730		
P for heterogeneity within main subsites [c]				0.12			0.05			0.41		
Calcium intake, per 300mg/day												
HR	0.96	0.95	0.88	0.94	0.87	0.83	0.84	0.90	0.92	0.91		
95% CI	0.87–1.05	0.88–1.02	0.78–1.01	0.89–0.99	0.76–1.00	0.77–0.90	0.79–0.90	0.79–1.04	0.84–1.00	0.85–0.98		
P_{trend}	0.33	0.18	0.06	0.02	0.047	<0.001	<0.001	0.15	0.06	0.02	0.06	0.19
P for heterogeneity within main subsites [c]				0.27			0.68			0.73		
Empirical dietary inflammatory pattern (EDIP) score, per 1 unit												
HR	1.54	1.16	1.01	1.27	1.24	1.52	1.45	1.21	1.13	1.19		
95% CI	1.06–2.23	0.84–1.61	0.61–1.69	1.02–1.58	0.72–2.12	1.13–2.05	1.12–1.88	0.69–2.15	0.79–1.60	0.89–1.61		
P_{trend}	0.02	0.36	0.96	0.03	0.44	0.006	0.005	0.51	0.51	0.24	0.50	0.71
P for heterogeneity within main subsites [c]				0.24			0.38			0.87		
Empirical dietary index for hyperinsulinemia (EDIH), per 1 unit												
HR	1.21	1.03	2.19	1.27	2.02	1.64	1.72	0.71	1.54	1.29		
95% CI	0.75–1.95	0.67–1.58	1.16–4.12	0.96–1.69	1.01–4.03	1.13–2.39	1.24–2.39	0.34–1.48	0.98–2.40	0.88–1.88		
P_{trend}	0.44	0.90	0.02	0.10	0.047	0.009	0.001	0.36	0.06	0.19	0.08	0.45
P for heterogeneity within main subsites [c]				0.11			0.60			0.05		

[a] Age- and cohort-stratified Cox proportional hazards model was used with adjustment for age (continuous), race, height (continuous), family history of colorectal cancer (yes or no), history of lower gastrointestinal endoscopic screening (yes or no), body mass index (continuous), pack-years of smoking (continuous), physical activity (continuous), alcohol intake (continuous), and regular aspirin use (yes or no).

[b] For dietary factors, test for trend was conducted using the median of each quintile as a continuous variable.

[c] Heterogeneity was tested by using meta-regression method with a subsite-specific random effect term as a nominal variable.

[d] Linear heterogeneity was tested by using meta-regression method with a subsite-specific random effect term as an ordinal variable.