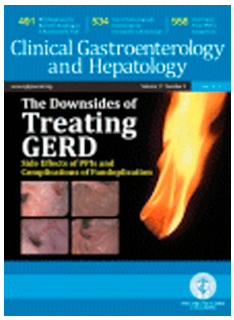
Accepted Manuscript



Heterogeneity of Colorectal Cancer Risk Factors by Anatomical Subsite in 10 European Countries: A Multinational Cohort Study

Neil Murphy, Heather A. Ward, Mazda Jenab, Joseph A. Rothwell, Marie-Christine Boutron-Ruault, Franck Carbonnel, Marina Kvaskoff, Rudolf Kaaks, Tilman Kühn, Heiner Boeing, Krasimira Aleksandrova, Elisabete Weiderpass, Guri Skeie, Kristin Benjaminsen Borch, Anne Tjønneland, Cecilie Kyrø, Kim Overvad, Christina C. Dahm, Paula Jakszyn, Maria-Jose Sánchez, Leire Gil, José M. Huerta, Aurelio Barricarte, J. Ramón Quirós, Kay-Tee Khaw, Nick Wareham, Kathryn E. Bradbury, Antonia Trichopoulou, Carlo La Vecchia, Anna Karakatsani, Domenico Palli, Sara Grioni, Rosario Tumino, Francesca Fasanelli, Salvatore Panico, Bas Bueno-de-Mesquita, Petra H. Peeters, Björn Gylling, Robin Myte, Karin Jirström, Jonna Berntsson, Xiaonan Xue, Elio Riboli, Amanda J. Cross, Marc J. Gunter

PII: S1542-3565(18)30756-0
DOI: 10.1016/j.cgh.2018.07.030
Reference: YJCGH 55976

To appear in: *Clinical Gastroenterology and Hepatology*
Accepted Date: 20 July 2018

Please cite this article as: Murphy N, Ward HA, Jenab M, Rothwell JA, Boutron-Ruault M-C, Carbonnel F, Kvaskoff M, Kaaks R, Kühn T, Boeing H, Aleksandrova K, Weiderpass E, Skeie G, Borch KB, Tjønneland A, Kyrø C, Overvad K, Dahm CC, Jakszyn P, Sánchez M-J, Gil L, Huerta JM, Barricarte A, Quirós JR, Khaw K-T, Wareham N, Bradbury KE, Trichopoulou A, La Vecchia C, Karakatsani A, Palli D, Grioni S, Tumino R, Fasanelli F, Panico S, Bueno-de-Mesquita B, Peeters PH, Gylling B, Myte R, Jirström K, Berntsson J, Xue X, Riboli E, Cross AJ, Gunter MJ, Heterogeneity of Colorectal Cancer Risk Factors by Anatomical Subsite in 10 European Countries: A Multinational Cohort Study, *Clinical Gastroenterology and Hepatology* (2018), doi: 10.1016/j.cgh.2018.07.030.

Heterogeneity of Colorectal Cancer Risk Factors by Anatomical Subsite in 10 European Countries: A Multinational Cohort Study

Neil Murphy[1,†]; Heather A Ward[2]; Mazda Jenab[1]; Joseph A Rothwell[1]; Marie-Christine Boutron-Ruault[3,4]; Franck Carbonnel[3,4,5]; Marina Kvaskoff[3,4]; Rudolf Kaaks[6]; Tilman Kühn[6]; Heiner Boeing[7]; Krasimira Aleksandrova[8]; Elisabete Weiderpass[9,10,11,12]; Guri Skeie[12]; Kristin Benjaminsen Borch[12]; Anne Tjønneland[13]; Cecilie Kyrø[13]; Kim Overvad[14]; Christina C Dahm[14]; Paula Jakszyn[15,16]; Maria-Jose Sánchez[17,18]; Leire Gil[19]; José M. Huerta[18,20]; Aurelio Barricarte[18,21,22]; J. Ramón Quirós[23]; Kay-Tee Khaw[24]; Nick Wareham[25]; Kathryn E. Bradbury[26]; Antonia Trichopoulou[27,28]; Carlo La Vecchia[27,29]; Anna Karakatsani[27,30]; Domenico Palli[31]; Sara Grioni[32]; Rosario Tumino[33]; Francesca Fasanelli[34]; Salvatore Panico[35]; Bas Bueno-de-Mesquita[2,36,37,38]; Petra H. Peeters[39]; Björn Gylling[40]; Robin Myte[41]; Karin Jirström[42]; Jonna Berntsson[42]; Xiaonan Xue[43]; Elio Riboli[2]; Amanda J Cross[2]; Marc J Gunter[1].

[†] Corresponding author

1. Section of Nutrition and Metabolism, International Agency for Research on Cancer, Lyon, France.
2. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.
3. CESP, Fac. de médecine - Univ. Paris-Sud, Fac. de médecine - UVSQ, INSERM, Université Paris-Saclay, 94805, Villejuif, France.
4. Gustave Roussy, F-94805, Villejuif, France.
5. Department of Gastroenterology, Bicêtre University Hospital, Assistance Publique des Hôpitaux de Paris, Le Kremlin Bicêtre, France.
6. German Cancer Research Center (DKFZ), Division of Cancer Epidemiology, Heidelberg, Germany.
7. Department of Epidemiology, German Institute of Human Nutrition, Potsdam-Rehbrücke, Germany.
8. Nutrition, Immunity and Metabolism Start-up Lab, Department of Epidemiology, Potsdam-Rehbrücke, Germany.
9. Department of Research, Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway.
10. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.
11. Genetic Epidemiology Group, Folkhälsan Research Center and Faculty of Medicine, University of Helsinki, Helsinki, Finland.
12. Department of Community Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway.
13. Danish Cancer Society Research Center, Copenhagen, Denmark.
14. Department of Public Health, Aarhus University, Aarhus, Denmark.
15. Unit of Nutrition, Environment and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of Oncology, L´Hospitallet de Llobregat, Barcelona, Spain.
16. Facultat de Ciències de la Salut Blanquerna, Universitat Ramon Llull, Barcelona, Spain.
17. Escuela Andaluza de Salud Pública. Instituto de Investigación Biosanitaria ibs.GRANADA. Hospitales Universitarios de Granada/Universidad de Granada, Granada, Spain.
18. CIBER de Epidemiología y Salud Pública (CIBERESP), Spain.

19. Public Health Division of Gipuzkoa, Research institute of BioDonostia, San Sebastian, Spain.
20. Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain.
21. Navarra Public Health Institute, Pamplona, Spain.
22. Navarra Institute for Health Research (IdiSNA) Pamplona, Spain.
23. Public Health Directorate, Asturias, Spain.
24. University of Cambridge School of Clinical Medicine, Clinical Gerontology Unit, Addenbrooke's Hospital, Cambridge, UK.
25. Medical Research Council Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK.
26. Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK.
27. Hellenic Health Foundation, Athens, Greece.
28. WHO Collaborating Center for Nutrition and Health, Unit of Nutritional Epidemiology and Nutrition in Public Health, Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Greece.
29. Department of Clinical Sciences and Community Health Università degli Studi di Milano, Italy.
30. Pulmonary Medicine Department, School of Medicine, National and Kapodistrian University of Athens, "ATTIKON" University Hospital, Haidari, Greece.
31. Cancer Risk Factors and Life-Style Epidemiology Unit, Cancer Research and Prevention Institute – ISPO, Florence, Italy.
32. Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy.
33. Cancer Registry and Histopathology Department, "Civic - M.P. Arezzo" Hospital, ASP Ragusa, Italy.
34. Unit of Cancer Epidemiology, Department of Medical Sciences, University of Turin, Turin, Italy.
35. Dipartimento di Medicina Clinica e Sperimentale, Federico II University, Naples, Italy.
36. Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), PO Box 1, 3720 BA Bilthoven, The Netherlands.
37. Department of Gastroenterology and Hepatology, University Medical Centre, Utrecht, The Netherlands.
38. Department of Social & Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia.
39. Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands.
40. Department of Medical Biosciences, Pathology, Umeå University, Umeå, Sweden.
41. Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden.
42. Division of Oncology and Pathology, Department of Clinical Sciences, Lund University, Lund, Sweden.
43. Albert Einstein College of Medicine, Bronx, New York, NY, U.S.

Corresponding author:

Neil Murphy, PhD

International Agency for Research on Cancer

150 cours Albert Thomas

69372 Lyon Cedex 08 France

Email: murphyn@iarc.fr

Tel: +33 4 72 73 85 08

Keywords: Colorectal cancer; proximal colon; distal colon; rectum; risk factors; anatomical subsite; heterogeneity

Short title: Colorectal Cancer Risk Factors by Tumor Site

Disclosure of Potential Conflicts of Interest: The authors declare no potential conflicts of interest.

Author Contributions: Conceived and designed the experiments: MJG NM. Analyzed the data: NM MJG. This article was written by NM and MJG with assistance from AJC, MJ, HAW, and taking into account the comments and suggestions of the co-authors. All co-authors had the opportunity to comment on the analysis and interpretation of the findings and approved the final version for publication. ER is the overall coordinator of the EPIC study, which he conceptualized, designed, and implemented in collaboration with the main investigators in the collaborating centers. All authors contributed to recruitment, data collection/acquisition, and are responsible for the ongoing follow-up and management of the EPIC cohort.

Abstract

Background and Aims

Colorectal cancer located at different anatomical subsites may have distinct etiologies and risk factors. Previous studies that have examined this hypothesis have yielded inconsistent results, possibly because most have been of insufficient size to identify heterogeneous associations with precision.

Methods

In the European Prospective Investigation into Cancer and Nutrition study, we used multivariable joint Cox proportional hazards models, which accounted for tumors at different anatomical sites (proximal colon, distal colon, and rectum) as competing risks, to examine the relationships between 14 established/suspected lifestyle, anthropometric, and reproductive/menstrual risk factors with colorectal cancer risk. Heterogeneity across sites was tested using Wald tests.

Results

After 14.9 years (median) follow-up of 521,330 men and women, 6,291 colorectal cancer cases occurred. Physical activity was inversely related to proximal colon and distal colon cancer, but not to rectal cancer (*P*-heterogeneity=0.03). Height was positively associated with proximal and distal colon cancer only, but not rectal cancer (*P*-heterogeneity=0.0001). For men, but not women, heterogeneous relationships were observed for body mass index (*P*-heterogeneity=0.008) and waist circumference (*P*-heterogeneity=0.03), with weaker positive associations found for rectal cancer, compared to proximal and distal colon cancer. Current smoking was associated with a greater risk of rectal and proximal colon cancer, but not distal colon cancer (*P*-heterogeneity=0.05). No heterogeneity by anatomical site was found for alcohol consumption, diabetes, nonsteroidal anti-inflammatory drug use, and reproductive/menstrual factors.

Conclusions

The physical activity, anthropometry, and smoking relationships with colorectal cancer risk differed by subsite, supporting the hypothesis that tumors in different anatomical regions may have distinct etiologies.

Keywords

Colorectal cancer; risk factors; anatomical subsite; heterogeneity

Introduction

Colorectal cancer (CRC) is one of the most frequently occurring malignancies worldwide. In 2012, 746,000 and 614,000 new cases were diagnosed globally in men (third most common cancer) and in women (second most common cancer), respectively[1]. Colorectal tumors at different anatomical sites have variable clinical characteristics[2]. In the proximal colon, tumors typically present at a later stage with a poorer prognosis than those in the distal colon and rectum[3, 4]. Women are more likely to develop cancers in the proximal colon, while in men cancers are more common in the distal colon region[5]. In addition, with advancing age, a greater proportion of colorectal tumors are located in the proximal colon, with a reduced proportion of rectal tumors[6].

Molecular heterogeneity has also been found for CRC tumors across anatomical sites. CpG island methylator phenotype (CIMP)-high, microsatellite instability (MSI)-high, and *PIK3CA* and *BRAF* mutations are most commonly found in the proximal colon region, with a linear decrease in frequency across the distal colon and rectum regions[7]. *KRAS* mutations have been found to be most common in the caecum region of the proximal colon, compared to other bowel regions[7]. *TP53* mutations are more frequent in tumors in the distal colon and rectum, compared to the proximal colon[8, 9].

CRC tumors at different anatomical locations may also have differential etiologies and risk factors[6, 8, 10, 11]. Previous studies that have examined this hypothesis have yielded inconsistent results, possibly because most have been of insufficient size to identify heterogeneous associations with precision. We, therefore, undertook a comprehensive investigation of how 14 established or suspected lifestyle, anthropometric, and reproductive and menstrual risk factors are associated with tumors located at the three main anatomical sites (proximal colon, distal colon, and rectum) in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, with >520,000 participants. The large number of incident CRC cases (>6,200) affords high statistical power to compare risk factor associations across tumor anatomical sites.

Methods

Study Population

EPIC is a multicenter prospective cohort of 521,448 participants mostly aged 35 years or above, who were recruited between 1992 and 2000, predominantly from the general population of 10 European countries (Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom)[12]. Written informed consent was provided by all study participants, and ethical approval for EPIC was provided by the International Agency for Research on Cancer and local participating centers. Participants with cancer diagnoses prior to recruitment (n=29,456); those in the highest and lowest 1% of the distribution for the ratio of energy intake to estimated energy requirement (n=9,573); and those with missing information on alcohol consumption and follow-up (n=6,259) were excluded from analyses. Additional exposure specific exclusions were applied when there was missing information for the risk factor of interest.

Exposures

The 14 CRC risk factors, all measured at recruitment, considered in the current analysis were: alcohol consumption (per 15 g/day); ever NSAID use (no, yes); physical activity index (inactive, moderately inactive, moderately active, active); prevalent diabetes (no, yes); smoking status (never, former, current); BMI (per 5 kg/m^2); height (per 10 cm); waist circumference (per 5 cm); waist-to-hip-ratio (per 0.05); and in women only, age at menarche (<12, 12-13, 14-15, ≥15 years); age at menopause (≤50, 51-52, 53-54, ≥55 years); ever OC use (never, ever); ever MHT use (never, ever); and duration of MHT use (never users, <2, 2-<5, 5-<8, ≥8 years). In secondary analyses, we investigated the relationships by anatomical subsite for alcohol consumption from wine (per 15 g/day), beer (per 15 g/day), and spirits liquors (per 3 g/day). Full details of measurements are detailed in the Supplementary Methods.

Follow-Up for Cancer Incidence and Vital Status

Cancer incidence was determined through record linkage with regional cancer registries or via a combination of methods, including the use of health insurance records, contacts with cancer and pathology registries, and active follow-up. CRC cases were defined using the Tenth Revision of the International Classification of Diseases (ICD-10) and the Second Revision of the International Classification of Diseases for Oncology (ICDO-2). Proximal colon cancer included those within the caecum, appendix, ascending colon, hepatic flexure, transverse colon, and splenic flexure (C18.0–18.5). Distal colon cancer included

those within the descending (C18.6) and sigmoid (C18.7) colon. Cancer of the rectum included cancer occurring at the recto-sigmoid junction (C19) and rectum (C20).

Statistical analysis

Hazard ratios (HRs) and the corresponding 95% confidence intervals (95%CIs) for the 14 risk factors and CRC were estimated using Cox proportional hazards models. Age was used as the time-scale in all models. Time at entry was age at recruitment. Exit time was age at whichever of the following came first: CRC diagnosis, death, or the last date at which follow-up was considered complete in each center. For the analyses by anatomical site, HRs and 95%CI were estimated using multivariable joint Cox proportional hazards model which accounted for tumors located at different anatomical sites as competing risks[13]. Heterogeneity across sites was tested using Wald tests. Full details on the statistical methods are in the Supplementary Methods and are detailed by Xue et al.[13]. Separate models were run for body size measurements and CRC for men and women due to *a priori* knowledge that the relationship differs by sex[14]. To determine whether the lifestyle risk factors and CRC relationships differed by sex, we included an interaction term for sex (multiplicative scale) in the model. The statistical significance of the cross-product terms was evaluated using the likelihood ratio test. Due to no heterogeneity being found by sex for smoking status (*P*-interaction=0.36), physical activity (*P*-interaction=0.71), alcohol consumption (*P*-interaction=0.45), diabetes (*P*-interaction=0.83), and NSAID use (*P*-interaction=0.34), men and women were analyzed together. Multivariable models were, where appropriate, mutually adjusted. We also conducted sensitivity analyses separating tumors located in the caecum (C18) into an additional anatomical site and examining heterogeneity in the relationships to each risk factor across four anatomical sites (caecum colon versus proximal colon versus distal colon versus rectum). Statistical tests used in the analysis were all two-sided and a *P*-value <0.05 was considered statistically significant.

Results

During a median follow-up of 14.9 years, 6,291 CRC cases occurred (2,718 in men and 3,573 in women). Of these, 1,877 were located in the proximal colon, 1,743 in the distal colon, and 2,094 in the rectum. Table 1 shows the characteristics of participants included in the analysis.

Alcohol consumption, prevalent diabetes, and smoking were associated with a greater risk of CRC, and ever NSAID use and physical activity were associated with a lower risk (Figure 1). For physical activity, compared to being inactive, the physically active group had a lower risk of developing CRC (HR=0.90, 95%CI: 0.82-0.98; *P*-trend=0.01). This inverse association was most evident for proximal colon cancers (HR=0.74, 95%CI: 0.63-0.87; *P*-trend=0.0004), while the estimates were not statistically significant for distal colon or rectal cancers (*P*-heterogeneity for proximal-distal-rectal=0.03). Smoking was associated with the development of CRC (current smokers versus never smokers, HR=1.19, 95%CI: 1.11-1.28; *P*-trend<0.0001). By anatomical site, heterogeneity was observed, with current smoking (versus never smokers) being associated with elevated risks of proximal colon (HR=1.19, 95%CI: 1.05-1.34) and rectal cancers (HR=1.27, 95%CI: 1.14-1.42), but not distal colon cancer (HR=1.08, 95%CI: 0.94-1.23) (*P*-heterogeneity across three sites=0.05; *P*-heterogeneity proximal and distal colon=0.04). Former smoking was associated with a greater risk of developing distal colon cancer (versus never smokers, HR=1.27, 95%CI: 1.13-1.43). Greater alcohol consumption was associated with elevated risk of CRC (HR per 15 g/day increment, HR=1.05, 95%CI: 1.03-1.07). Although the test for heterogeneity was not statistically significant (*P*-heterogeneity=0.15 for proximal-distal-rectal), positive associations were found for distal colon and rectal cancer, but not for proximal colon cancer. No heterogeneity was observed for tumors located at different anatomical subsites for alcohol from wine, beer, and spirits/liquors when analyzed separately (all *P*-heterogeneities>0.05) (Table S1). Prevalent diabetes at baseline (yes versus no) was associated with higher CRC risk (HR=1.28, 95%CI: 1.12-1.47), with similar positive relationships found across anatomical sites (*P*-heterogeneity>0.70), although the association for rectal cancer was not statistically significant. Ever use of NSAIDs was associated with a lower CRC risk (versus never use, HR=0.85, 95%CI: 0.74-0.99), with no heterogeneity observed for tumors located at different anatomical sites (all *P*-heterogeneity>0.30).

For men and women, higher BMI, height, waist circumference, and waist-to-hip ratio were all associated with greater risk of CRC (Figure 2). For men, the positive relationship for BMI was weaker for rectal cancer (HR per 5 kg/m^2, HR=1.10, 95%CI: 1.01-1.20), compared to proximal colon (HR per 5 kg/m^2, HR=1.31, 95%CI: 1.18-1.47) and distal colon cancers (HR per 5 kg/m^2, HR=1.32, 95%CI: 1.20-1.45) (*P*-heterogeneity=0.008), but no heterogeneity was found between tumors in the proximal and distal colon (*P*-heterogeneity=0.94). Also in men, the positive waist circumference association was weaker

for tumors located in the rectum (HR per 5 cm, HR=1.06, 95%CI: 1.03-1.09), than for tumors in the proximal (HR per 5 cm, HR=1.11, 95%CI: 1.07-1.16) and distal colon (HR per 5 cm, HR=1.12, 95%CI: 1.08-1.16) (*P*-heterogeneity=0.03), but no heterogeneity was found across the colon (proximal versus distal *P*-heterogeneity=0.78). The positive association between waist-to-hip ratio and CRC for men and women was consistent across all anatomical sites (all *P*-heterogeneities>0.60). For men and women, height was not associated with rectal cancer (men HR per 10 cm, HR=0.97, 95%CI: 0.88-1.06; women HR per 10 cm, HR=0.92, 95%CI: 0.83-1.03), but was positively related to both proximal colon and distal colon cancers (*P*-heterogeneity=0.0001 for men and *P*-heterogeneity<0.0001 for women). The association of height with colon cancer did not differ between proximal and distal colon in men (*P*-heterogeneity=0.24), but there was some suggestion of heterogeneity for women (*P*-heterogeneity=0.05), with a stronger positive association observed for proximal colon cancer (HR per 10 cm, HR=1.30, 95%CI: 1.17-1.43) than for distal colon cancer (HR per 10 cm, HR=1.11, 95%CI: 0.99-1.25). For women, no heterogeneity by subsite was observed for the other anthropometric measurements, with similar strength associations found for BMI, waist circumference, and waist-to-hip ratio across tumors at the three anatomical sites (all *P*-heterogeneities>0.05).

Ever MHT use versus never use was associated with a lower risk of CRC (HR=0.90, 95%CI: 0.83-0.97), with no evidence of heterogeneity across subsites (*P*-heterogeneities>0.16) (Figure 3). Duration of MHT use was inversely associated with CRC risk (P-trend=0.01), with no heterogeneity found by anatomical site (*P*-heterogeneity>0.05). Age at menarche and ever OC use was not associated with CRC and no heterogeneity was observed across anatomical sites (*P*-heterogeneity>0.05). Older age (≥55 years) versus younger age at menopause (≤50 years) was associated with elevated CRC risk (HR=1.20, 95%CI: 1.03-1.38), with similar relationships observed by anatomical site (*P*-heterogeneity>0.40).

When tumors located in the caecum were considered as an additional subsite end-point, a similar pattern of heterogeneous relationships was considered across the four subsites (caecum colon, proximal colon, distal colon, and rectum) (Tables S2 to S4).

Discussion

In this multi-country prospective study, we found heterogeneous relationships by tumor site for physical activity, smoking, and anthropometric measurements. Low levels of physical activity and greater height and BMI were primarily associated with an increased risk of distal or proximal colon cancer, with weaker or null relationships found for rectal cancer. Current smoking was associated with an increased risk of proximal colon and rectal cancer, while no heterogeneity by anatomical site was found for alcohol consumption, prevalent diabetes, NSAIDs use, and, in women, reproductive and menstrual factors.

For overall CRC, we observed the expected pattern of risk factor associations. Greater adiposity and height were associated with elevated CRC risk, as were higher alcohol consumption, smoking, prevalent diabetes, and later age at menopause. Conversely, being physically active, and use of NSAIDs and MHT were associated with lower risk of developing CRC. Our analysis benefited from the large number of incident CRC cases which accrued during the longer follow-up period, which allowed well-powered analyses for the 14 risk factors by tumor anatomical site. Recently, a similar analysis of CRC risk factors by anatomical site was undertaken in a large UK cohort, with no heterogeneity found for the considered risk factors by tumor anatomical site[15]; however, that study included only women, so it is uncertain whether the findings are generalizable to men[15]. Previous studies which have investigated heterogeneity in the association between major risk factors and colorectal anatomical subsites in men and women had smaller numbers of cases compared to our analysis, and may have been constrained by insufficient statistical power to identify weak-to-moderate strength heterogeneous associations[16, 17]. In the current study, which included men and women, we observed heterogeneous relationships between several risk factors and tumors across different anatomical sites.

We found that greater physical activity was similarly related to lower risks of developing tumors in the proximal and distal colon regions, findings consistent with other large prospective studies[15, 17], and a meta-analysis of 21 studies[18]. Physical activity was not, however, related to rectal cancer risk, a result inconsistent with a recent participant-level pooled analysis which reported an inverse relationship between physical activity and rectal cancer incidence[19], but in accordance with a joint Nurses' Health Study and Health Professionals Follow-up Study analysis[10]. The biological mechanisms through which physical activity potentially lowers colon cancer risk, but not rectal cancer risk, are uncertain. Being physically active is associated with less weight gain and body fatness[20], and therefore has a

beneficial effect on CRC risk[21]. However, in our study, we found that greater BMI and waist circumference were risk factors for colon and, albeit more weakly, for rectal cancer. Greater physical activity has also been associated with lower insulin levels and beneficial effects on inflammatory pathways and dyslipidemia, including lowering levels of circulating triglycerides[22-24]. Previous meta-analyses suggest that C-peptide (a marker of insulin secretion), C-reactive protein (a nonspecific marker of systemic inflammation), and triglycerides are positively associated with colon, but not with rectal cancer[25-28]. This suggests that any beneficial effects of physical exercise on insulin (or correlated metabolic markers), inflammatory, and lipid pathways would be more likely to influence tumors in the colon, and not in the rectum, potentially explaining the null result we observed for physical activity with rectal cancer.

Our finding that higher BMI was more strongly related to greater CRC risk among men than among women is in accordance with a large body of epidemiological evidence[21, 29, 30]. We observed heterogeneous relationships for anthropometric measurements by anatomical site, particularly for men. For BMI, the positive relationship found among men was weaker for rectal cancer compared to tumors in the colon. A meta-analysis of prospective studies also observed that, for men, greater BMI was more weakly associated with rectal cancer (relative risk [RR] per 5 kg/m^2 unit increase in BMI=1.12, 95%CI: 1.09-1.16) than with colon cancer (RR per 5 kg/m^2 unit increase in BMI=1.30, 95%CI: 1.25-1.35)[21]. A moderately weaker positive relationship was found for waist circumference and rectal cancer in men compared to colonic subsites, yet for waist-to-hip ratio no heterogeneity by anatomical site was observed. For men and women, height was associated with colon cancer, but not with rectal cancer. This null result for rectal cancer is inconsistent with other large prospective cohort studies and a meta-analysis which found a positive association for height and rectal cancer[31, 32]. Additionally, positive relationships of similar magnitude were found for both colon and rectal cancer in a Mendelian randomization analysis[33].

Current smoking was related to an elevated risk of proximal colon and rectal cancers, but not distal colon cancer. A similar pattern of results for smoking history as found in the Nurses' Health Study, with 40 pack-years of smoking (versus none) only being positively associated with proximal colon (HR=1.31, 95%CI: 1.16-1.48) and rectal cancer (HR=1.27, 95%CI: 1.05-1.53), but not distal colon cancer (HR=1.04, 95%CI: 0.88-1.23)[17]. MSI-high, *BRAF* mutation-positive, and CIMP-positive tumors, are more common in the proximal colon

region compared to the distal colon[7], and have been positively associated with cigarette smoking[11]. However, these molecular characteristics are even less common for malignant tumors in the rectum, the subsite for which we observed the strongest positive relationship with smoking. Additionally, a positive relationship was observed for former smokers and distal colon cancer which is inconsistent with these molecular characteristics explaining these findings.

The current investigation is the largest study to date to comprehensively investigate the relationships between CRC risk factor by anatomical site in both men and women. Limitations of our analysis were that all of the considered risk factors were measured once at baseline, and due to multiple known or suspected CRC risk factors being simultaneously investigated, some of our results could have been chance findings. Finally, our study would have been enhanced with information on tumor molecular features.

In conclusion, heterogeneous relationships across tumors located in the proximal colon, distal colon, and rectum were observed for physical activity, anthropometric measurements, and smoking. These results, taken together with the varying biological and molecular features of tumors located across the colorectum, indicate that tumors in different anatomical regions may have distinct etiologies.

Figures and legends

Figure 1. Multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) for colorectal cancer incidence for both sexes combined in relation to lifestyle factors, by anatomical site.

For alcohol consumption, physical activity, and smoking status: Multivariable models–Cox regression using age as the underlying time variable and stratified by sex, center, and age at recruitment. Models mutually adjusted, and additionally adjusted for body mass index, height, education level, ever use of menopausal hormone therapy, and intakes of alcohol, red and processed meats, calcium, and fiber. For ever nonsteroidal anti-inflammatory drug (NSAID) use and prevalent diabetes: Multivariable models–Cox regression using age as the underlying time variable and stratified by sex, center, and age at recruitment adjusted for body mass index, height, physical activity; smoking status and intensity; education level; ever use of menopausal hormone therapy; and intakes of alcohol, red and processed meats, calcium, and fiber. †Nonsteroidal anti-inflammatory drug (NSAID) use information only available from six centers-Cambridge, Utrecht, Heidelberg, Potsdam, Aarhus, Copenhagen.

Figure 2. Multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) for colorectal cancer incidence for both sexes combined in relation to anthropometric measures, by anatomical site.

Multivariable models only–Cox regression using age as the underlying time variable and stratified by center and age at recruitment, and adjusted for physical activity, smoking status and intensity, education level, ever use of menopausal hormone therapy, and intakes of alcohol, red and processed meats, calcium, and fiber. Multivariable model for height was further adjusted for body mass index. Multivariable models for body mass index, waist circumference, and waist-to-hip ratio were further adjusted for height.

Figure 3. Multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) for colorectal cancer incidence in relation to reproductive and menstrual factors among women, by anatomical site.

Multivariable models only–Cox regression using age as the underlying time variable and stratified by center and age at recruitment, and adjusted for body mass index, height, physical activity, smoking status and intensity, education level, ever use of menopausal hormone therapy, and intakes of alcohol, red and processed meats, calcium and fiber. MHT-menopausal hormone therapy.

Funding

The coordination of EPIC is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); Nordic Centre of Excellence programme on Food, Nutrition and Health (Norway); Health Research Fund (FIS), PI13/00061 to Granada; PI13/01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia (no. 6236) and Navarra, ISCIII RETIC (RD06/0020) (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom).

For information on how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php

Table 1. Characteristics of participants at recruitment‡

	Both sexes				
	Non-cases	**Colorectal cancer cases**	**Colon proximal cancer cases**	**Colon distal cancer cases**	**Rectal cancer cases**
N	469,869	6,291	1,877	1,743	2,094
Women (%)	70.3	56.8	64.4	56.0	50.7
Age at recruitment (years)	51.2 (9.9)	57.3 (7.9)	58.2 (7.9)	56.9 (7.5)	56.6 (7.7)
Alcohol consumption (g/day)	11.6 (16.8)	15.0 (20.2)	12.6 (18.4)	15.4 (20.5)	16.5 (21.4)
Smoking status					
Never (%)	49.1	40.7	43.6	40.4	38.4
Current (%)	22.4	24.1	22.8	22.3	26.0
Ever nonsteroidal anti-inflammatory drug (NSAID) use					
Yes (%)	8.2	8.5	8.2	9.4	8.3
Physical activity					
Inactive (%)	20.9	24.9	27.9	25.0	21.8
Active (%)	17.9	18.4	15.6	18.7	21.4
Prevalent diabetes					
Yes (%)	2.8	4.4	4.5	4.6	3.8
Body mass index (kg/m^2)					
Men	26.5 (3.6)	27.2 (3.8)	27.3 (4.0)	27.5 (3.8)	26.9 (3.6)
Women	25.4 (4.6)	26.1 (4.6)	25.9 (4.5)	26.3 (4.7)	26.0 (4.5)
Height (cm)					
Men	174.7 (7.4)	174.4 (7.1)	175.2 (7.1)	174.5 (7.3)	174.2 (7.0)
Women	161.8 (6.8)	161.8 (6.6)	162.3 (6.2)	161.7 (6.6)	161.5 (6.4)
Waist circumference (cm)					
Men	94.6 (10.2)	97.4 (10.2)	97.6 (10.4)	98.2 (10.5)	96.8 (9.9)
Women	80.2 (11.5)	82.6 (11.7)	82.6 (11.5)	83.1 (12.1)	82.0 (11.7)
Waist-to-hip ratio					
Men	0.94 (0.1)	0.96 (0.1)	0.95 (0.1)	0.96 (0.1)	0.96 (0.1)

Women	0.79 (0.1)	0.81 (0.1)	0.81 (0.1)	0.81 (0.1)	0.80 (0.1)
Age at menarche (years)	13.1 (1.5)	13.2 (1.6)	13.2 (1.6)	13.2 (1.6)	13.2 (1.5)
Age at menopause (years)	48.6 (5.0)	49.0 (5.0)	49.0 (5.0)	49.0 (4.8)	49.2 (5.1)
Ever oral contraceptive (OC) use					
Yes (%)	58.8	47.5	45.3	48.2	51.9
Ever menopausal hormone therapy (MHT) use					
Yes (%)	25.9	31.1	32.8	29.5	30.9
Education					
Longer education (including university)	24.2	19.0	19.1	18.4	18.8
Red and processed meat intake (g/day)	74.7 (51.0)	83.0 (52.7)	78.8 (51.3)	82.7 (52.3)	87.2 (53.5)
Calcium intake (mg/day)	994.8 (409.4)	985.0 (398.5)	994.1 (392.6)	970.4 (393.6)	984.2 (401.3)
Fibre intake (g/day)	22.8 (7.7)	22.6 (7.7)	22.5 (7.6)	22.5 (7.9)	22.8 (7.5)

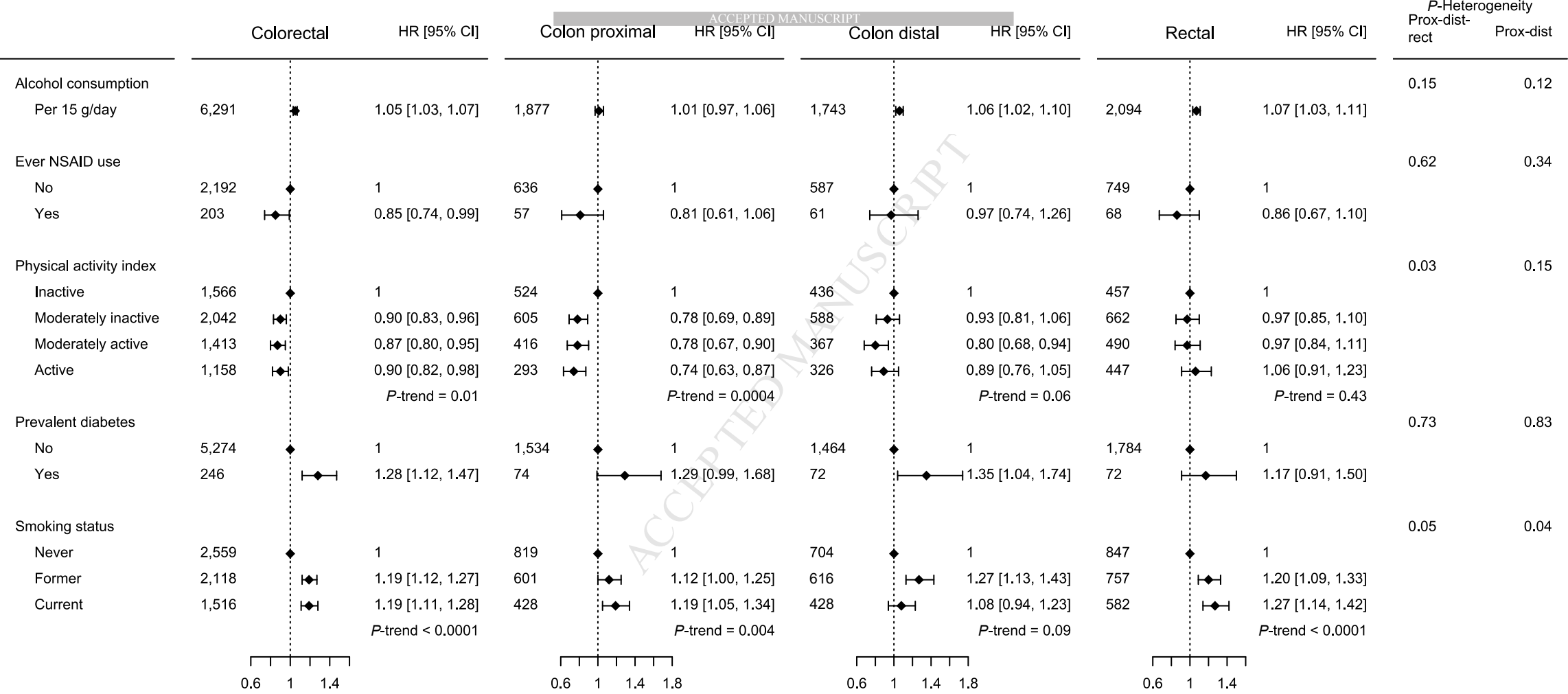
Mean and standard deviation unless stated otherwise.

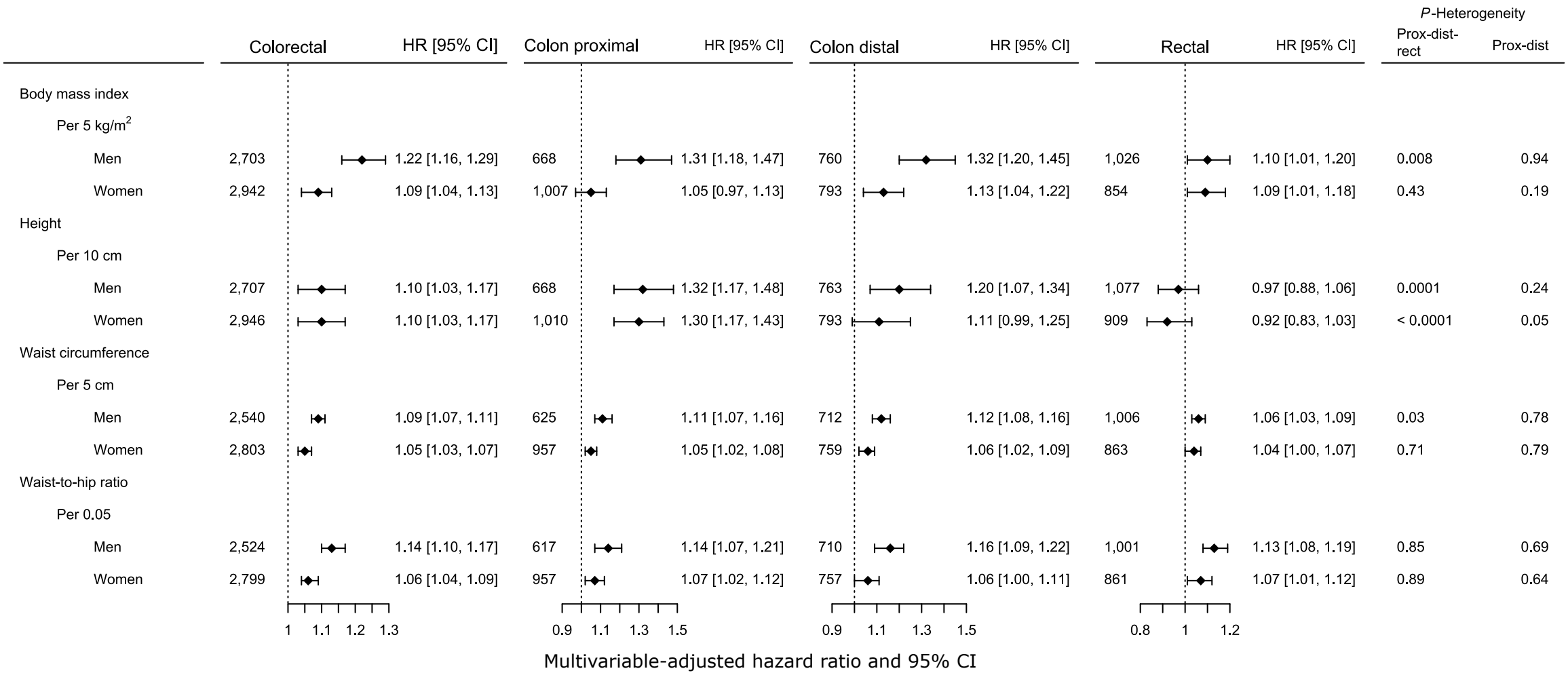
[‡]Based on participant numbers in the alcohol consumption models.

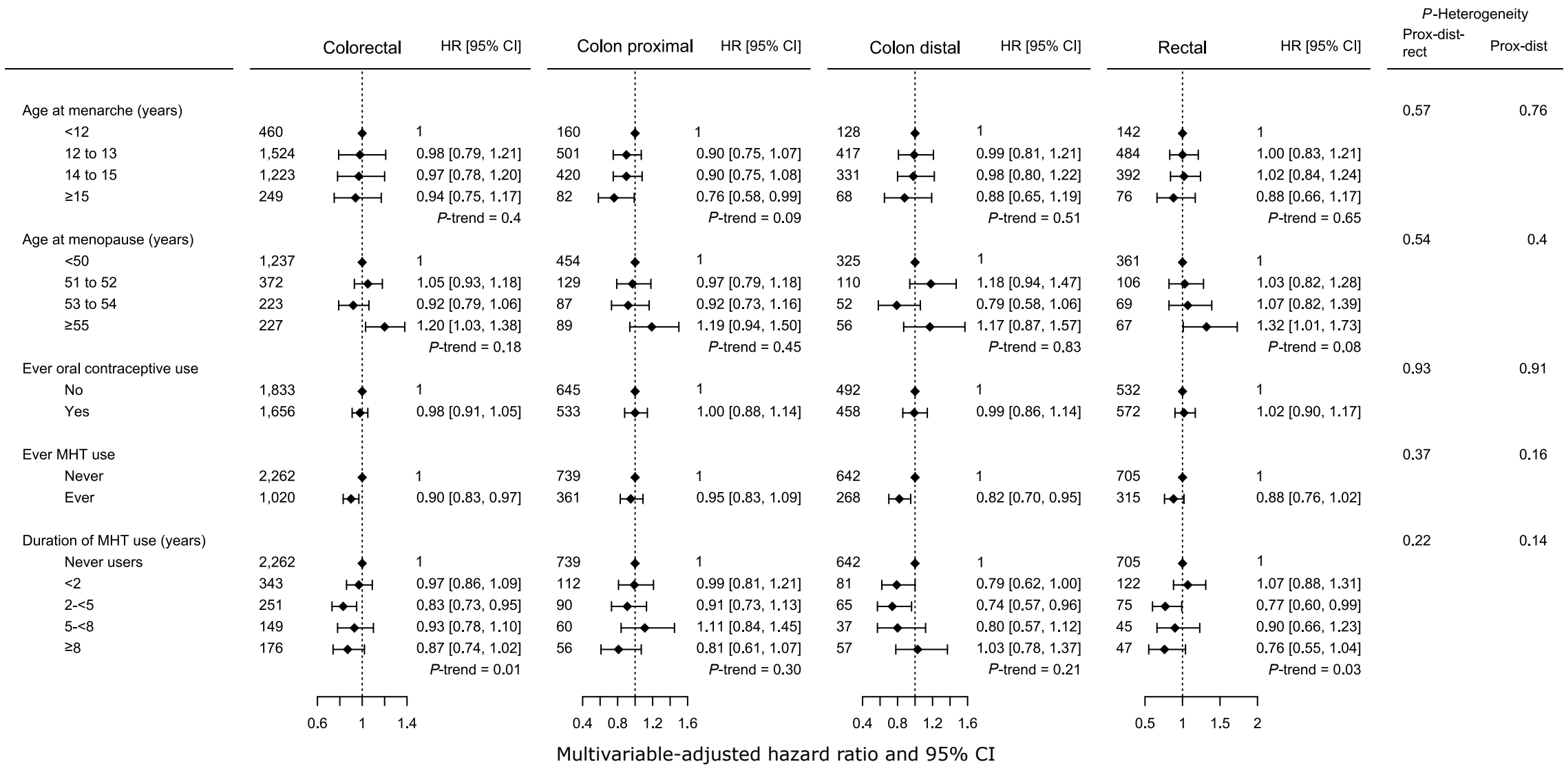
References

1. IARC-WHO. GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012, 2015.
2. Kocarnik JM, Shiovitz S, Phipps AI. Molecular phenotypes of colorectal cancer and potential clinical applications. Gastroenterology Report 2015;3:269-276.
3. Wong R. Proximal Tumors Are Associated with Greater Mortality in Colon Cancer. Journal of General Internal Medicine 2010;25:1157-1163.
4. Phipps AI, Lindor NM, Jenkins MA, et al. Colon and Rectal Cancer Survival by Tumor Location and Microsatellite Instability: The Colon Cancer Family Registry. Diseases of the colon and rectum 2013;56:937-944.
5. Carethers JM. One Colon Lumen but Two Organs. Gastroenterology 2011;141:411-412.
6. Siegel R, DeSantis C, Jemal A. Colorectal cancer statistics, 2014. CA: A Cancer Journal for Clinicians 2014;64:104-117.
7. 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;61:847-854.
8. Iacopetta B. Are there two sides to colorectal cancer? International Journal of Cancer 2002;101:403-408.
9. Russo A, Bazan V, Iacopetta B, et al. The TP53 Colorectal Cancer International Collaborative Study on the Prognostic and Predictive Significance of p53 Mutation: Influence of Tumor Site, Type of Mutation, and Adjuvant Treatment. Journal of Clinical Oncology 2005;23:7518-7528.
10. Wei EK, Giovannucci E, Wu K, et al. Comparison of risk factors for colon and rectal cancer. International Journal of Cancer 2004;108:433-442.
11. Limsui D, Vierkant RA, Tillmans LS, et al. Cigarette Smoking and Colorectal Cancer Risk by Molecularly Defined Subtypes. Journal of the National Cancer Institute 2010;102:1012-1022.
12. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. International Journal of Epidemiology 1997;26:S6-14.
13. Xue X, Kim MY, Gaudet MM, et al. A Comparison of the Polytomous Logistic Regression and Joint Cox Proportional Hazards Models for Evaluating Multiple Disease Subtypes in Prospective Cohort Studies. Cancer Epidemiology Biomarkers & Prevention 2013;22:275-285.
14. Pischon T, Lahmann PH, Boeing H, et al. Body Size and Risk of Colon and Rectal Cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). Journal of the National Cancer Institute 2006;98:920-931.
15. Burón Pust A, Alison R, Blanks R, et al. Heterogeneity of colorectal cancer risk by tumour characteristics: Large prospective study of UK women. International Journal of Cancer 2017;140:1082-1090.
16. Shin A, Joo J, Bak J, et al. Site-Specific Risk Factors for Colorectal Cancer in a Korean Population. PLOS ONE 2011;6:e23196.
17. Wei EK, Colditz GA, Giovannucci EL, et al. A Comprehensive Model of Colorectal Cancer by Risk Factor Status and Subsite Using Data From the Nurses' Health Study. American Journal of Epidemiology 2017;185:224-237.
18. Boyle T, Keegel T, Bull F, et al. Physical activity and risks of proximal and distal colon cancers: a systematic review and meta-analysis. J Natl Cancer Inst 2012;104:1548-61.

19. Moore SC, Lee IM, Weiderpass E, et al. Association of Leisure-Time Physical Activity With Risk of 26 Types of Cancer in 1.44 Million Adults. JAMA Intern Med 2016;176:816-25.
20. Lee I, Djoussé L, Sesso HD, et al. Physical activity and weight gain prevention. JAMA 2010;303:1173-1179.
21. Larsson SC, Wolk A. Obesity and colon and rectal cancer risk: a meta-analysis of prospective studies. The American Journal of Clinical Nutrition 2007;86:556-565.
22. Helmerhorst HJF, Wijndaele K, Brage S, et al. Objectively Measured Sedentary Time May Predict Insulin Resistance Independent of Moderate- and Vigorous-Intensity Physical Activity. Diabetes 2009;58:1776-1779.
23. Woods JA, Vieira VJ, Keylock KT. Exercise, Inflammation, and Innate Immunity. Neurologic Clinics 2006;24:585-599.
24. Szapary PO, Bloedon LT, Foster GD. Physical activity and its effects on lipids. Current Cardiology Reports 2003;5:488-493.
25. Zhou B, Shu B, Yang J, et al. C-reactive protein, interleukin-6 and the risk of colorectal cancer: a meta-analysis. Cancer Causes & Control 2014;25:1397-1405.
26. Tsilidis KK, Branchini C, Guallar E, et al. C-reactive protein and colorectal cancer risk: A systematic review of prospective studies. International Journal of Cancer 2008;123:1133-1140.
27. Chen L, Li L, Wang Y, et al. Circulating C-peptide level is a predictive factor for colorectal neoplasia: evidence from the meta-analysis of prospective studies. Cancer Causes & Control 2013;24:1837-1847.
28. Yao X, Tian Z. Dyslipidemia and colorectal cancer risk: a meta-analysis of prospective studies. Cancer Causes & Control 2015;26:257-268.
29. Keimling M, Renehan AG, Behrens G, et al. Comparison of Associations of Body Mass Index, Abdominal Adiposity, and Risk of Colorectal Cancer in a Large Prospective Cohort Study. Cancer Epidemiology Biomarkers & Prevention 2013;22:1383-1394.
30. Moghaddam AA, Woodward M, Huxley R. Obesity and Risk of Colorectal Cancer: A Meta-analysis of 31 Studies with 70,000 Events. Cancer Epidemiology Biomarkers & Prevention 2007;16:2533-2547.
31. Green J, Cairns BJ, Casabonne D, et al. Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. The Lancet Oncology 2011;12:785-794.
32. WCRF-AICR. Diet, nutrition, physical activity and colorectal cancer. Continuous Update Project. http://www.wcrf.org/sites/default/files/CUP%20Colorectal%20Report_2017_Digital.pdf, 2017.
33. Thrift AP, Gong J, Peters U, et al. Mendelian randomization study of height and risk of colorectal cancer. Int J Epidemiol 2015;44:662-72.




Colorectal
HR [95% CI]
Colon proximal
HR [95% CI]
Colon distal
HR [95% CI]
Rectal
HR [95% CI]
P-Heterogeneity
Prox-dist-rect
Prox-dist
Body mass index
Per 5 kg/m2
Men 2,703 1.22 [1.16, 1.29] 668 1.31 [1.18, 1.47] 760 1.32 [1.20, 1.45] 1,026 1.10 [1.01, 1.20] 0.008 0.94
Women 2,942 1.09 [1.04, 1.13] 1,007 1.05 [0.97, 1.13] 793 1.13 [1.04, 1.22] 854 1.09 [1.01, 1.18] 0.43 0.19
Height
Per 10 cm
Men 2,707 1.10 [1.03, 1.17] 668 1.32 [1.17, 1.48] 763 1.20 [1.07, 1.34] 1,077 0.97 [0.88, 1.06] 0.0001 0.24
Women 2,946 1.10 [1.03, 1.17] 1,010 1.30 [1.17, 1.43] 793 1.11 [0.99, 1.25] 909 0.92 [0.83, 1.03] < 0.0001 0.05
Waist circumference
Per 5 cm
Men 2,540 1.09 [1.07, 1.11] 625 1.11 [1.07, 1.16] 712 1.12 [1.08, 1.16] 1,006 1.06 [1.03, 1.09] 0.03 0.78
Women 2,803 1.05 [1.03, 1.07] 957 1.05 [1.02, 1.08] 759 1.06 [1.02, 1.09] 863 1.04 [1.00, 1.07] 0.71 0.79
Waist-to-hip ratio
Per 0.05
Men 2,524 1.14 [1.10, 1.17] 617 1.14 [1.07, 1.21] 710 1.16 [1.09, 1.22] 1,001 1.13 [1.08, 1.19] 0.85 0.69
Women 2,799 1.06 [1.04, 1.09] 957 1.07 [1.02, 1.12] 757 1.06 [1.00, 1.11] 861 1.07 [1.01, 1.12] 0.89 0.64
1 1.1 1.2 1.3
0.9 1.1 1.3 1.5
0.9 1.1 1.3 1.5
0.8 1 1.2
Multivariable-adjusted hazard ratio and 95% CI


Colorectal HR [95% CI]
Colon proximal HR [95% CI]
Colon distal HR [95% CI]
Rectal HR [95% CI]
P-Heterogeneity
Prox-dist-rect
Prox-dist
Age at menarche (years) 0.57 0.76
<12 460 1 160 1 128 1 142 1
12 to 13 1,524 0.98 [0.79, 1.21] 501 0.90 [0.75, 1.07] 417 0.99 [0.81, 1.21] 484 1.00 [0.83, 1.21]
14 to 15 1,223 0.97 [0.78, 1.20] 420 0.90 [0.75, 1.08] 331 0.98 [0.80, 1.22] 392 1.02 [0.84, 1.24]
≥15 249 0.94 [0.75, 1.17] 82 0.76 [0.58, 0.99] 68 0.88 [0.65, 1.19] 76 0.88 [0.66, 1.17]
P-trend = 0.4 P-trend = 0.09 P-trend = 0.51 P-trend = 0.65
Age at menopause (years) 0.54 0.4
<50 1,237 1 454 1 325 1 361 1
51 to 52 372 1.05 [0.93, 1.18] 129 0.97 [0.79, 1.18] 110 1.18 [0.94, 1.47] 106 1.03 [0.82, 1.28]
53 to 54 223 0.92 [0.79, 1.06] 87 0.92 [0.73, 1.16] 52 0.79 [0.58, 1.06] 69 1.07 [0.82, 1.39]
≥55 227 1.20 [1.03, 1.38] 89 1.19 [0.94, 1.50] 56 1.17 [0.87, 1.57] 67 1.32 [1.01, 1.73]
P-trend = 0.18 P-trend = 0.45 P-trend = 0.83 P-trend = 0.08
Ever oral contraceptive use 0.93 0.91
No 1,833 1 645 1 492 1 532 1
Yes 1,656 0.98 [0.91, 1.05] 533 1.00 [0.88, 1.14] 458 0.99 [0.86, 1.14] 572 1.02 [0.90, 1.17]
Ever MHT use 0.37 0.16
Never 2,262 1 739 1 642 1 705 1
Ever 1,020 0.90 [0.83, 0.97] 361 0.95 [0.83, 1.09] 268 0.82 [0.70, 0.95] 315 0.88 [0.76, 1.02]
Duration of MHT use (years) 0.22 0.14
Never users 2,262 1 739 1 642 1 705 1
<2 343 0.97 [0.86, 1.09] 112 0.99 [0.81, 1.21] 81 0.79 [0.62, 1.00] 122 1.07 [0.88, 1.31]
2-<5 251 0.83 [0.73, 0.95] 90 0.91 [0.73, 1.13] 65 0.74 [0.57, 0.96] 75 0.77 [0.60, 0.99]
5-<8 149 0.93 [0.78, 1.10] 60 1.11 [0.84, 1.45] 37 0.80 [0.57, 1.12] 45 0.90 [0.66, 1.23]
≥8 176 0.87 [0.74, 1.02] 56 0.81 [0.61, 1.07] 57 1.03 [0.78, 1.37] 47 0.76 [0.55, 1.04]
P-trend = 0.01 P-trend = 0.30 P-trend = 0.21 P-trend = 0.03
0.6 1 1.4
0.4 0.8 1.2 1.6
0.4 0.8 1.2 1.6
0.5 1 1.5 2
Multivariable-adjusted hazard ratio and 95% CI

What You Need to Know

Background

- Previous research indicates that colorectal tumours located at different anatomical sites have distinct clinical and molecular characteristics.
- It has also been hypothesized that colorectal cancer at different anatomical locations may have differential aetiologies and risk factors.
- Previous epidemiological studies may have been underpowered to detect heterogeneous relationships by anatomical site.

Findings

- This was the largest study to date to comprehensively investigate the relationships between colorectal cancer risk factors by anatomical site in both men and women, with >520,000 participants from 10 European countries included, and >6,200 incident colorectal cancer cases.
- We found heterogeneous relationships across tumours located in the proximal colon, distal colon, and rectum for physical activity levels, anthropometric measurements, and smoking.

Implications

- These results highlight the importance of separating the colorectum into distinct entities with separate aetiologies.
- Variability in the carcinogenic processes at different sites of the large-bowel may explain the complex risk factor-colorectal cancer relationships.

Methods

Exposures

The 14 colorectal cancer risk factors, all measured at recruitment, considered in the current analysis were: alcohol consumption (per 15 g/day increment); ever NSAID use (no, yes); physical activity index (inactive, moderately inactive, moderately active, active); prevalent diabetes (no, yes); smoking status (never, former, current); BMI (per 5 kg/m^2 increment); height (per 10 cm increment); waist circumference (per 5 cm increment); waist-to-hip-ratio (per 0.05 increment); and in women only, age at menarche (<12, 12 to 13, 14 to 15, ≥15 years); age at menopause (≤50, 51 to 52, 53 to 54, ≥55 years); ever OC use (never, ever); ever MHT use (never, ever); and duration of MHT use (never users, <2, 2-<5, 5-<8, ≥8 years). In secondary analyses, we investigated the relationships by anatomical subsite for alcohol consumption from wine (per 15 g/day increment), beer (per 15 g/day increment), and spirits liquors (per 3 g/day increment).

With participants not wearing shoes, weight was measured to the nearest 0.1 kg and height was measured—dependent on the study center—to the nearest 0.1, 0.5, or 1.0 cm. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared (kg/m^2). Waist circumference was measured either at the narrowest torso circumference or at the midpoint between the lower ribs and iliac crest. Hip circumference was measured at the widest circumference (France; Italy; Spain; Bilthoven, The Netherlands; Greece; Malmö, Sweden) or over the buttocks (the United Kingdom; Utrecht, The Netherlands; Germany; Denmark). Waist-to-hip ratio was calculated by dividing waist circumference by hip circumference. Standardized lifestyle and personal history questionnaires were collected at recruitment[1, 2], before disease onset or diagnosis. Information on cigarette smoking habits included baseline smoking status (never, former, or current smoker). Overall physical activity (the sum/total of occupational physical activity and leisure time physical activity) was assessed from three questions referring to the past year and an index was derived by allocating individuals to four categories of overall activity (inactive, moderately inactive, moderately active and active)[3]. Information was collected on education, diabetes prevalence, oral contraceptive (OC) use, menopausal hormone therapy (MHT) use, age at menarche, age at menopause, and, in six centers (Cambridge, UK; Utrecht, The Netherlands; Heidelberg and Potsdam, Germany; Aarhus and Copenhagen, Denmark),

NSAID use (including aspirin). Diet over the previous 12 months was assessed at recruitment using validated country/centre-specific dietary questionnaires[1, 2]. Alcohol consumption at recruitment was calculated from the number of standard glasses of beer, wine, cider, sweet liquor, distilled spirits or fortified wines consumed per day/week reported during the 12 months prior to recruitment.

Follow-Up for Cancer Incidence and Vital Status

Cancer incidence was determined through record linkage with regional cancer registries (Denmark, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom) or via a combination of methods, including the use of health insurance records, contacts with cancer and pathology registries, and active follow-up through participants and their next of kin (France, Germany, and Greece). Colorectal cancer cases were defined using the Tenth Revision of the International Classification of Diseases (ICD-10) and the Second Revision of the International Classification of Diseases for Oncology (ICDO-2). Proximal colon cancer included those within the caecum, appendix, ascending colon, hepatic flexure, transverse colon, and splenic flexure (C18.0–18.5). Distal colon cancer included those within the descending (C18.6) and sigmoid (C18.7) colon. Cancer of the rectum included cancer occurring at the recto-sigmoid junction (C19) and rectum (C20).

Statistical analysis

Hazard ratios (HRs) and the corresponding 95% confidence intervals (95%CIs) for the 14 risk factors and CRC were estimated using Cox proportional hazards models. Age was used as the time-scale in all models. Time at entry was age at recruitment. Exit time was age at whichever of the following came first: colorectal cancer diagnosis, death, or the last date at which follow-up was considered complete in each center. Possible non-proportionality was assessed using an analysis of Schoenfeld residuals[4], with no evidence of non-proportionality being detected. For the analyses by anatomical site, HRs and 95%CI were estimated using multivariable joint Cox proportional hazards model which accounted for tumors located at different anatomical sites as competing risks[5]. The heterogeneity in baseline risk of colorectal cancer subsites was addressed by stratified Cox models where each subsite was allowed to have its own baseline hazard function; the heterogeneity in association with risk factors across

subsites was assessed by including an interaction term between each risk factor and the indicators of colorectal cancer subsites and testing the statistical significance of the interaction terms. As a robust variance was used to address the competing risk between colorectal cancer subsites, a log-likelihood ratio test was no longer valid. We, therefore, used a global Wald-test based on the robust variance estimates obtained from a "sandwich" type of estimator. Full details on the statistical method are in the Supplementary Methods and are detailed by Xue et al.[5]

References

1. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. International Journal of Epidemiology 1997;26:S6-14.
2. Riboli E, Hunt KJ, Slimani N, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutrition 2002;5:1113-1124.
3. Consortium TI. Validity of a short questionnaire to assess physical activity in 10 European countries. European Journal of Epidemiology 2012;27:15-25.
4. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69:239-241.
5. Xue X, Kim MY, Gaudet MM, et al. A Comparison of the Polytomous Logistic Regression and Joint Cox Proportional Hazards Models for Evaluating Multiple Disease Subtypes in Prospective Cohort Studies. Cancer Epidemiology Biomarkers & Prevention 2013;22:275-285.

Table S1. Multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) for colorectal cancer incidence for both sexes combined in relation to alcohol intake (overall and by source), by anatomical site

	Both sexes							
	Colorectal cancer		Colon proximal		Colon distal		Rectal	
	N cases	Multivariable	N cases	Multivariable	N cases	Multivariable	N cases	Multivariable
Alcohol								
Per 15g/day	6291	1.05 (1.03-1.07)	1877	1.01 (0.97-1.06)	1743	1.06 (1.02-1.10)	2094	1.07 (1.03-1.11)
P-Heterogeneity proximal-distal-rectal				0.15				
P-Heterogeneity proximal-distal				0.12				
Alcohol from wine								
Per 15g/day	6291	1.03 (0.99-1.06)	1877	1.00 (0.93-1.07)	1743	1.05 (1.00-1.11)	2094	1.04 (0.99-1.09)
P-Heterogeneity proximal-distal-rectal				0.46				
P-Heterogeneity proximal-distal				0.22				
Alcohol from beer								
Per 15g/day	6291	1.09 (1.05-1.13)	1877	1.03 (0.94-1.12)	1743	1.10 (1.03-1.17)	2094	1.11 (1.06-1.16)
P-Heterogeneity proximal-distal-rectal				0.29				
P-Heterogeneity proximal-distal				0.21				
Alcohol from spirits/liquors								
Per 3g/day	6291	1.01 (1.00-1.03)	1877	1.00 (0.97-1.04)	1743	1.00 (0.96-1.03)	2094	1.02 (1.00-1.05)
P-Heterogeneity proximal-distal-rectal				0.27				
P-Heterogeneity proximal-distal				0.80				

Multivariable models only – Cox regression using age as the underlying time variable and stratified by sex, center, and age at recruitment. Models adjusted for body mass index, height, physical activity index, smoking status and intensity, education level attained, ever use of menopausal hormone therapy, and intakes of red and processed meats, dietary calcium, and fiber.

Table S2. Multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) for colorectal cancer incidence for both sexes combined in relation to lifestyle factors, by tumors in the colon caecum, colon proximal, colon distal, and rectum.

	Both sexes							
	Colon caecum		Colon proximal		Colon distal		Rectal	
	N cases	Multivariable	N cases	Multivariable	N cases	Multivariable	N cases	Multivariable
Alcohol consumption								
Per 15g/day	720	1.00 (0.92-1.09)	1198	1.03 (0.97-1.08)	1743	1.06 (1.02-1.10)	2211	1.07 (1.04-1.11)
P-Heterogeneity caecum-proximal-distal-rectal				0.33				
Ever nonsteroidal anti-inflammatory drug (NSAID) use[†]								
No	257	1	587	1	587	1	802	1
Yes	28	0.91 (0.61-1.35)	61	0.73 (0.50-1.05)	61	0.97 (0.74-1.26)	73	0.86 (0.67-1.09)
P-Heterogeneity caecum-proximal-distal-rectal				0.67				
Physical activity index								
Inactive	196	1	344	1	436	1	457	1
Moderately inactive	231	0.88 (0.72-1.09)	383	0.72 (0.62-0.84)	588	0.93 (0.81-1.06)	662	0.97 (0.86-1.10)
Moderately active	156	0.88 (0.69-1.13)	270	0.73 (0.60-0.87)	367	0.80 (0.69-0.94)	490	0.99 (0.87-1.15)
Active	113	0.83 (0.64-1.08)	186	0.68 (0.56-0.83)	326	0.90 (0.76-1.05)	447	1.07 (0.93-1.24)
P-trend		0.18		0.0003		0.06		0.29
P-Heterogeneity caecum-proximal-distal-rectal				0.02				
Prevalent diabetes								
No	559	1	1012	1	1464	1	1784	1
Yes	23	1.29 (0.84-2.00)	54	1.33 (0.97-1.82)	72	1.34 (1.04-1.74)	72	1.21 (0.95-1.54)
P-Heterogeneity caecum-proximal-distal-rectal				0.94				
Smoking status								
Never	320	1	509	1	704	1	847	1
Former	233	1.07 (0.89-1.27)	385	1.18 (1.03-1.36)	616	1.27 (1.13-1.43)	757	1.20 (1.09-1.33)

Current	151	1.12 (0.91-1.38)	289	1.25 (1.08-1.46)	388	1.08 (0.94-1.23)	582	1.27 (1.14-1.42)
P-trend		0.27		0.0017		0.09		<0.0001
P-Heterogeneity caecum-proximal-distal-rectal				0.13				

For alcohol consumption, physical activity index, and smoking status: Multivariable models only – Cox regression using age as the underlying time variable and stratified by sex, center, and age at recruitment. Models mutually adjusted, and additionally adjusted for body mass index, height, education level attained, ever use of menopausal hormone therapy, and intakes of alcohol, red and processed meats, dietary calcium, and fiber.

For ever nonsteroidal anti-inflammatory drug (NSAID) use and prevalent diabetes: Multivariable models only – Cox regression using age as the underlying time variable and stratified by sex, center, and age at recruitment adjusted for body mass index, height, physical activity index; smoking status and intensity; education level attained; ever use of menopausal hormone therapy; and intakes of alcohol, red and processed meats, dietary calcium, and fiber.

† Nonsteroidal anti-inflammatory drug (NSAID) use information only available from six centers (Cambridge, UK; Utrecht, The Netherlands; Heidelberg and Potsdam, Germany; Aarhus and Copenhagen, Denmark).

Table S3. Multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) for colorectal cancer incidence for both sexes combined in relation to anthropometric measures, by tumors in the colon caecum, colon proximal, colon distal, and rectum.

	Colon caecum		Colon proximal		Colon distal		Rectal	
	N cases	**Multivariable**	**N cases**	**Multivariable**	**N cases**	**Multivariable**	**N cases**	**Multivariable**
BMI								
Men								
Per 5 kg/m^2	250	1.41 (1.19-1.68)	437	1.26 (1.09-1.45)	760	1.32 (1.20-1.45)	1076	1.11 (1.02-1.03)
P-Heterogeneity caecum-proximal-distal-rectal				0.01				
Women								
Per 5 kg/m^2	405	1.06 (0.94-1.19)	624	1.06 (0.97-1.16)	793	1.13 (1.04-1.22)	854	1.08 (1.01-1.16)
P-Heterogeneity caecum-proximal-distal-rectal				0.72				
Height								
Men								
Per 10 cm	250	1.43 (1.18-1.75)	437	1.22 (1.06-1.42)	763	1.20 (1.07-1.34)	1077	0.95 (0.86-1.04)
P-Heterogeneity caecum-proximal-distal-rectal				<0.0001				
Women								
Per 10 cm	407	1.30 (1.11-1.52)	625	1.26 (1.11-1.45)	793	1.10 (0.99-1.25)	909	0.92 (0.83-1.03)
P-Heterogeneity caecum-proximal-distal-rectal				0.0003				
Waist circumference								
Men								
Per 5 cm	236	1.13 (1.06-1.20)	409	1.10 (1.05-1.16)	712	1.12 (1.08-1.17)	1006	1.06 (1.03-1.09)
P-Heterogeneity caecum-proximal-distal-rectal				0.05				
Women								
Per 5 cm	389	1.04 (0.99-1.09)	591	1.06 (1.02-1.10)	759	1.06 (1.02-1.09)	863	1.04 (1.00-1.07)
P-Heterogeneity caecum-proximal-distal-rectal				0.78				

Waist-to-hip ratio								
Men								
Per 0.05	233	1.14 (1.04-1.25)	404	1.12 (1.04-1.21)	710	1.16 (1.09-1.22)	1001	1.13 (1.08-1.19)
P-Heterogeneity caecum-proximal-distal-rectal				0.93				
Women								
Per 0.05	389	1.04 (0.96-1.13)	591	1.10 (1.04-1.16)	757	1.06 (1.00-1.11)	861	1.07 (1.01-1.12)
P-Heterogeneity caecum-proximal-distal-rectal				0.62				

Multivariable models only – Cox regression using age as the underlying time variable and stratified by center and age at recruitment, and adjusted for physical activity index, smoking status and intensity, education level attained, ever use of menopausal hormone therapy, and intakes of alcohol, red and processed meats, dietary calcium, and fiber. Multivariable model for height was further adjusted for body mass index. Multivariable models for body mass index, waist circumference, and waist-to-hip ratio were further adjusted for height.

Table S4. Multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) for colorectal cancer incidence among women in relation to reproductive and menstrual characteristics, by tumors in the colon caecum, colon proximal, colon distal, and rectum.

	Colon caecum		Colon proximal		Colon distal		Rectal	
	N cases	**Multivariable**	**N cases**	**Multivariable**	**N cases**	**Multivariable**	**N cases**	**Multivariable**
Age at menarche (years)								
<12	13	1	23	1	28	1	22	1
12 to 13	144	0.92 (0.52-1.63)	205	0.75 (0.48-1.15)	276	0.89 (0.61-1.32)	325	1.30 (0.84-2.00)
14 to 15	203	0.77 (0.44-1.36)	348	0.75 (0.49-1.15)	460	0.92 (0.63-1.36)	545	1.34 (0.87-2.06)
≥15	92	0.78 (0.43-1.41)	157	0.79 (0.51-1.23)	180	0.88 (0.59-1.32)	202	1.21 (0.78-1.89)
P-trend		0.1372		0.9997		0.7919		0.9427
P-Heterogeneity caecum-proximal-distal-rectal				0.54				
Age at menopause (years)								
≤50	172	1	287	1	325	1	361	1
51 to 52	53	1.06 (0.77-1.44)	78	0.93 (0.72-1.20)	110	1.18 (0.94-1.47)	106	1.03 (0.82-1.28)
53 to 54	31	0.84 (0.57-1.23)	59	1.02 (0.77-1.35)	52	0.79 (0.58-1.06)	69	1.07 (0.82-1.39)
≥55	46	1.52 (1.10-2.12)	48	1.05 (0.76-1.43)	56	1.17 (0.87-1.57)	67	1.32 (1.01-1.73)
P-trend		0.1281		0.8442		0.8376		0.0794
P-Heterogeneity caecum-proximal-distal-rectal				0.44				
Ever oral contraceptive use								
No	276	1	380	1	492	1	532	1
Yes	179	0.82 (0.66-1.01)	365	1.14 (0.97-1.34)	458	0.99 (0.86-1.14)	572	1.02 (0.90-1.17)
P-Heterogeneity caecum-proximal-distal-rectal				0.11				
Ever menopausal hormone therapy use								
Never		1		1		1		1
Ever		0.90 (0.83-0.97)		0.95 (0.83-1.09)		0.82 (0.70-0.95)		0.88 (0.76-1.02)

P-Heterogeneity caecum-proximal-distal-rectal				0.37				
Duration of menopausal hormone therapy use (years)								
Never users	288	1	467	1	642	1	705	1
<2	46	1.08 (0.78-1.50)	68	0.93 (0.72-1.21)	81	0.79 (0.62-1.00)	122	1.07 (0.88-1.31)
2-<5	34	0.94 (0.65-1.35)	60	0.92 (0.70-1.22)	65	0.74 (0.57-0.96)	75	0.77 (0.60-0.99)
5-<8	21	1.00 (0.63-1.60)	39	1.12 (0.80-1.57)	37	0.79 (0.56-1.12)	45	0.90 (0.66-1.23)
≥8	23	0.76 (0.49-1.19)	34	0.82 (0.57-1.18)	57	1.03 (0.78-1.37)	47	0.76 (0.55-1.04)
P-trend		0.34		0.46		0.21		0.03
P-Heterogeneity caecum-proximal-distal-rectal				0.46				

Multivariable models only – Cox regression using age as the underlying time variable and stratified by center and age at recruitment, and adjusted for body mass index, height, physical activity index, smoking status and intensity, education level attained, ever use of menopausal hormone therapy, and intakes of alcohol, red and processed meats, dietary calcium and fiber.