A prospective nutrient wide association study for risk of colorectal cancer

Nikos Papadimitriou[1,2*], Emmanouil Bouras[1*], Piet A van den Brandt[3], David C Muller[4], Areti Papadopoulou[1], Alicia K Heath[4], Elena Critselis[5], Marc J Gunter[2], Paolo Vineis[4], Pietro Ferrari[2], Elisabete Weiderpass[2], Heiner Boeing[6], Nadia Bastide[7], Melissa A Merritt[8], David S Lopez[9,10], Manuela M Bergmann[11], Aurora Perez-Cornago[12], Matthias Schulze[13,14], Guri Skeie[15], Bernard Srour[16], Anne Kirstine Eriksen[17], Stina Boden[18], Ingegerd Johansson[19], Therese Haugdahl Nøst[15], Marko Lukic[15], Fulvio Ricceri[20,21], Ulrika Ericson[22], José María Huerta[23,24], Christina C Dahm[25], Claudia Agnoli[26], Pilar Amiano Exezarreta[27,28], Anne Tjønneland[29,30], Aurelio Barricarte Gurrea[31], Bas Bueno-de-Mesquita[32], Eva Ardanaz[28,33,34], Jonna Berntsson[35], Maria-Jose Sánchez[28,36,37,38], Rosario Tumino[39], Salvatore Panico[40], Verena Katzke[41], Paula Jakszyn[42,43], Giovanna Masala[44], Jeroen Derksen[45], **other EPIC co-authors**, Amanda J Cross[4], Elio Riboli[4], Ioanna Tzoulaki[1,4], Konstantinos K Tsilidis[1,4]

*These authors had equal contribution

[1]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece

[2]International Agency for Research on Cancer (IARC), Lyon, France

[3]Department of Epidemiology, GROW School for Oncology and Developmental Biology, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, Netherlands

[4]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK

[5]Biomedical Research Foundation of the Academy of Athens, Athens, Greece

[6]Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Germany

[7]Association BRCA France, Montpellier, France

[8]University of Hawaii Cancer Center, Honolulu, Hawaii.

[9]Department of Preventive Medicine and Population Health – UTMB School of Medicine, Galveston, TX, USA.

[10]Division of Urology, UTHealth McGovern Medical School, Houston, TX, USA.

[11]German Institute of Human Nutrition Potsdam-Rehbrücke, Germany

[12]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

[13]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany

[14]Institute of Nutrition Science, University of Potsdam, Nuthetal, Germany

[15]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø-The Arctic University of Norway, Tromsø, Norway
[16]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[17]Danish Cancer Society Research Center, Diet, Genes and Environment, Copenhagen, Denmark
[18]Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden
[19]Department of Odontology, Umeå University, Sweden
[20]Department of Clinical and Biological Sciences, University of Turin, Italy
[21]Unit of Epidemiology, Regional Health Service ASL TO3, Grugliasco (TO), Italy
[22]Department of Clinical Sciences in Malmö, Lund University, Malmö, Sweden
[23]Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain
[24]CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[25]Department of Public Health, Aarhus University, Aarhus, Denmark
[26]Epidemiology and Prevention Unit Fondazione IRCCS Istituto Nazionale dei Tumori di Milano Via Venezian, Milano, Italy
[27]Public Health Division of Gipuzkoa, BioDonostia Research Institute, Donostia-San Sebastian, Spain
[28]CIBER Epidemiology and Public Health CIBERESP, Madrid, Spain
[29]Danish Cancer Society Research Center, Diet, Genes and Environment, Copenhagen, Denmark
[30]Department of Public Health, University of Copenhagen, Copenhagen, Denmark
[31]Instituto de Salud Pública de Navarra, Navarra, Spain
[32]Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
[33]Navarra Public Health Institute, Pamplona, Spain
[34]IdiSNA, Navarra Institute for Health Research, Pamplona, Spain
[35]Department of Clinical Sciences Lund, Oncology and Pathology, Lund University, Lund, Sweden
[36]Escuela Andaluza de Salud Pública (EASP), Granada, Spain
[37]Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain
[38]Department of Preventive Medicine and Public Health, University of Granada, Granada, Spain.
[39]Cancer Registry and Histopathology Department, Provincial Health Authority (ASP), Ragusa, Italy
[40]DIPARTIMENTO DI MEDICINA CLINICA E CHIRURGIA FEDERICO II UNIVERSITY, NAPLES, ITALY
[41]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[42]Unit of Nutrition and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of Oncology (ICO-IDIBELL), Barcelona, Spain
[43]Blanquerna School of Health Sciences, Ramon Llull University, Barcelona, Spain
[44]Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network - ISPRO, Florence, ITALY

[45]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Corresponding author:
Dr Konstantinos K Tsilidis, Department of Epidemiology and Biostatistics, Imperial College London, St Mary's Campus, London, W2 1PG, United Kingdom, Tel.: +44 (0) 2075942623, E-mail: k.tsilidis@imperial.ac.uk

Novelty and Impact (max 75 words)

Evidence regarding the association of dietary exposures with colorectal cancer (CRC) risk is not consistent. We conducted a nutrient-wide association study (NWAS) in EPIC to systematically evaluate the associations between various food and nutrient intakes with CRC risk and replicated in an independent cohort, the NLCS. Results confirmed previously reported associations for alcohol, dairy and calcium and suggested a lower CRC risk following higher intakes of phosphorus, magnesium, potassium, riboflavin, beta-carotene and total protein.

Research Article

Abstract (max 250 words, unstructured)

The association of 92 food and nutrient intakes with colorectal cancer (CRC) risk was assessed using a nutrient-wide association approach in 386,792 participants, 5,069 of whom developed incident CRC, of the European Prospective Investigation into Cancer and Nutrition (EPIC). Correction for multiple comparisons was performed using the false discovery rate, and emerging associations were examined in the Netherlands Cohort Study (NLCS). Multiplicative gene-nutrient interactions were also tested in EPIC based on known CRC-associated loci. In EPIC, alcohol, liquor/spirits, wine, beer/cider, soft drinks, and pork were positively associated with CRC, whereas milk, cheese, calcium, phosphorus, magnesium, potassium, riboflavin, vitamin B6, beta-carotene, fruit, fibre, non-white bread, banana, and total protein intakes were inversely associated. Of these 20 associations, 13 were replicated in NLCS, for which a meta-analysis was performed, namely alcohol (summary HR per 1 SD increment in intake: 1.07; 95%CI: 1.04-1.09), liquor/spirits (1.04; 1.02-1.06), wine (1.04; 1.02-1.07), beer/cider (1.06; 1.04-1.08), milk (0.95; 0.93-0.98), cheese (0.96; 0.94-0.99), calcium (0.93; 0.90-0.95), phosphorus (0.92; 0.90-0.95), magnesium (0.95; 0.92-0.98), potassium (0.96; 0.94-0.99), riboflavin (0.94; 0.92-0.97), beta-carotene (0.96; 0.93-0.98), and total protein (0.94; 0.92-0.97). None of the gene-nutrient interactions were significant after adjustment for multiple comparisons. Our findings confirm a positive association for alcohol and an inverse association for dairy products and calcium with CRC risk, and also suggest a lower risk at higher dietary intakes of phosphorus, magnesium, potassium, riboflavin, beta-carotene and total protein.

Keywords: Diet; cohort study; colorectal cancer; epidemiology; nutrition

Abbreviations: BMI: Body mass index; CI: Confidence interval; CRC: Colorectal cancer; EPIC: European Prospective Investigation into Cancer and Nutrition; FDR: False discovery rate; GWAS: Genome-wide association study; HR: Hazard ratio; NLCS: the Netherlands Cohort Study; NOS: Non-specified; NWAS: Nutrient-wide association study; SNP: Single nucleotide polymorphism; WCRF: World Cancer Research Fund; WGS: Whole-genome sequencing.

Introduction

Colorectal cancer (CRC) is the third most common type of cancer worldwide with over 1.8 million new cases and over 800,000 deaths in 2018[1]. The incidence rates are higher in high income countries, but there has been a recent large increase in the rates in low- and middle-income countries potentially due to the "westernization" of these societies[1]. Several aspects of the Western lifestyle such as obesity and lack of physical activity are well-established risk factors of CRC[2, 3], but evidence regarding diet, and in particular the association of specific foods and nutrients with CRC is not consistent, with a few exceptions[4]. The World Cancer Research Fund (WCRF) third Expert Report identified strong evidence that consuming processed meat, red meat, and alcohol increases risk of CRC, whereas consumption of whole-grains, foods containing dietary fibre, and dairy products lowers CRC risk[4]. Associations for other foods and nutrients and CRC risk exist, but are inconsistent and currently provide limited evidence according to WCRF[4].

The aim of this study was to systematically examine the associations between a wide set of dietary factors and risk of CRC in the European Prospective Investigation into Cancer and Nutrition (EPIC) and the Netherlands Cohort Study (NLCS), by conducting a nutrient-wide association study (NWAS)[5-7]. The NWAS takes an analogous strategy to that of a genome-wide association study (GWAS) by separately estimating associations for each food and nutrient, using adjustments for multiple comparisons, and replicating promising associations in an independent study.

Materials and Methods

Study populations

EPIC is a large European multicentre prospective cohort that consists of 521,324 participants, mostly aged between 35 and 70 years, recruited between 1992 and 2000 from 23 centres across 10 European countries, namely Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom[8]. Out of the 491,992 participants with complete data on length of follow-up and without a cancer diagnosis before the baseline assessment, 6,259 were excluded because they did not complete the lifestyle or dietary questionnaires at baseline, 9,573 participants were excluded due to extreme values (top or bottom 1%) of the energy intake to energy requirement ratio, and 64,671 were further excluded due to missing values in any of the covariates of interest (diabetes history: 38,972; level of education:16,931; smoking status: 9,678; physical activity: 8,824). Data from Greece were also excluded from the current analysis, leaving 386,792 participants (71% women) in the final analytical sample. All participants gave written informed consent while approval for the study was obtained from the ethical review boards of the International Agency for Research on Cancer (IARC) and all local institutions in the participating countries.

NLCS is a prospective cohort study of 120,852 participants, aged between 55 and 69 years and recruited in 1986 from 204 computerised population registries across the Netherlands[9]. The NLCS used a case-cohort approach for efficiency reasons, whereby a subcohort of 5,000 participants was selected at random immediately after baseline[9]. Of the 5,000 participants, 3,893 were included in the current analysis after excluding 226 with prevalent cancer at recruitment, 690 with incomplete or inconsistent dietary data, and 191 participants with missing data on confounders. NLCS was approved by the institutional review boards of the Nederlandse Organisatie voor Toegepast Natuurwetenschappehlijk Onderzoek (TNO) Quality of Life research institute (Zeist, Netherlands) and Maastricht University (Maastricht, Netherlands).

Assessment of dietary factors

In EPIC, consumption of foods over the last 12 months was assessed at baseline using validated country-specific food questionnaires[8]. In most countries and centres the questionnaires were self-administered apart from Ragusa (Italy) and Spain, where interviewers were used. In Malmö (Sweden), a food record was used for cooked meals and a food frequency questionnaire was used for breakfast and foods consumed between the main meals. The EPIC Nutrient Database (ENDB) was used to calculate standardized nutrient intakes[10]. In total, 92 dietary factors (63 foods and 29 nutrients) that were available in at least 8 out of the 9 countries, were included in the current analysis.

In NLCS, information on dietary intake over the preceding 12 months was assessed at baseline using a semi-quantitative 150-item food frequency questionnaire, which has been validated and tested for reproducibility[11, 12]. The Dutch food composition table was used for the conversion of the data obtained from the food questionnaires to nutrient intakes[13].

Identification of colorectal cancer cases

In EPIC, incident CRC cases were identified by record linkage with population-based cancer registries in Denmark, Italy, The Netherlands, Norway, Spain, Sweden and UK, or a combination of registries, insurance records and active follow up of the study participants or their relatives in France, Germany and Naples (Italy). The 10th Revision of the International Classification of Diseases (ICD-10) and the second revision of the International Classification of Diseases for Oncology (ICD-O) were used to determine CRC cases (codes C18-C20).

In NLCS, incident CRC cases were identified by record linkage to the Netherlands Cancer Registry and the Dutch National Pathology Registry record[14]. CRC cases were classified according to ICD-O3 (codes C18-C20).

In addition to overall CRC, we also examined associations for the following subsites: proximal colon (C18.0–18.5), distal colon (C18.6–18.7), and rectum (C19-C20).

Statistical analyses

In EPIC, separate Cox proportional hazards regression models with age as the time scale were used to investigate the associations between each of the dietary factors with CRC risk. Age at recruitment was set as the age at entry. Age at exit was defined either as the age at cancer diagnosis or the age at death or age at the last follow-up, whichever occurred first. In NLCS, given the case-cohort design, Prentice weighted Cox proportional hazards regression models with robust standard error estimation were implemented[15]. In both EPIC and NLCS the proportionality of the hazard ratios (HR) was verified by examining the slope of the Schoenfeld residuals, and no violations were found. Intakes of foods and nutrients were adjusted for energy intake using the residual method and standardized prior to modelling [16]. All of the models were adjusted for: total energy intake (kcal, continuous); smoking status (never, former, current); body mass index (BMI, kg/m^2, <20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥ 35); physical activity [EPIC: Cambridge index (inactive, moderately inactive, moderately active, active), NLCS: non-occupational physical activity (≤30, >30-60, >60-90, >90 min/day)]; diabetes history (no, yes); level of education (none/primary school, technical/professional school, secondary school, longer education) and family history of CRC (no, yes; in NLCS only), and reflect associations per one standard deviation increase in daily consumption. Additionally, all models were further stratified by sex, age at recruitment (5-year intervals), and in EPIC also by centre in order to control for centre-specific differences like questionnaire design and follow-up procedures[17].

To account for multiple comparisons, the false discovery rate (FDR) was estimated for each association analysed using the sequential p-value approach proposed by Benjamini and Hochberg[18]. The dietary factors with an FDR less than 0.05 were subsequently selected for replication in NLCS, and fixed effects meta-analysis was performed to combine the results from the two cohorts when heterogeneity was low or moderate (p-value for heterogeneity>0.1 and/or $I^2 \leq 50\%$). To further investigate the robustness of the associations that were replicated in NLCS, a mutual adjustment model was used.

Separate analyses for the FDR-significant dietary exposures were conducted in men and women and also by anatomical subsite of CRC. For the FDR-significant foods or nutrients in EPIC, the pairwise partial correlation coefficients were quantified, adjusting for age, sex and centre, using Spearman's rho (ρ). Additionally, the impact of follow-up duration in the association of red and processed meat with CRC risk was investigated. All analyses were performed using R[19].

Gene-Nutrient interactions

Potential multiplicative gene-nutrient interactions in EPIC were systematically investigated, between the food components that met the FDR threshold and known CRC-associated genetic variants from GWAS[20]. Of the approximately 100 GWAS-identified SNPs associated with CRC, data for 73 SNPs or their proxies were available for 3,361 participants. Nutrients were included in the interaction analyses as standardized continuous variables and the same covariates as in the NWAS Cox proportional hazards regression models were used. P-values were adjusted for multiple comparisons using the Bonferroni correction based on the number of independent tests, with a corrected p-value threshold at 3.4×10^{-5}.

Results

Study characteristics

After a mean follow up of 14.1 years, a total of 5,069 (56.8% in women) incident malignant CRC cases were identified among the 386,792 included EPIC participants, of which 3,143 were identified as colon (1,495 proximal; 1,435 distal; 213 unspecified CRC) and 1,715 as rectal cancers. In NLCS, 3,765 cases (42.8% female) with incident and microscopically confirmed CRC were included in the present analysis, of which 2,612 were colon (1,348 proximal; 1,187 distal) and 801 were rectal cancers.

The main baseline characteristics of the study participants are shown in ***Table 1***. In EPIC, approximately 30% of the participants were men, and 47% were overweight or obese. About 50% of the participants were never smokers, and 47% were physically active. More than half of the NLCS subcohort participants were male (54%), one third (33%) were never smokers and 47% were overweight or obese, while 48% spent more than 60 minutes per day on non-occupational physical activities.

NWAS in EPIC

Of the 92 dietary factors that were examined in EPIC, 20 were associated with CRC risk (FDR<0.05) (***Figure 1***, ***Supplementary Table 1***). Higher intakes of alcohol (HR per 1 SD increment in intake/day = 1.07, 95%CI:1.04-1.10), liquor/spirits (1.03, 1.01-1.06), wine (1.05, 1.02-1.08), beer/cider (1.07, 1.04-1.09), soft drinks (1.04, 1.02-1.07), and pork (1.06, 1.03-1.09) were positively associated with CRC, whereas higher milk (0.96, 0.93-0.99), cheese (0.95, 0.92-0.99), calcium (0.92, 0.89-0.95), phosphorus (0.92, 0.89-0.94), magnesium (0.95, 0.91-0.99), potassium (0.95, 0.92-0.98), riboflavin (0.94, 0.91-0.98), vitamin B6 (0.95, 0.92-0.99), beta-carotene (0.95, 0.92-0.98), fruit (0.96, 0.92-0.99), fibre (0.93, 0.90-0.96), non-white bread (0.93, 0.90-0.97), banana (0.96, 0.93-0.99), and total protein (0.94, 0.91-0.97) intakes were associated with a lower CRC risk.

After conducting the analysis by tumour subsite, evidence of heterogeneity between colon and rectal cancer was observed for intakes of magnesium, potassium, vitamin B6 and banana (p-value for heterogeneity < 0.1), with associations being inverse for colon cancer and null for rectal cancer (***Supplementary Table 2***). Regarding proximal versus distal colon subsites, only total alcohol and wine had heterogeneous results (p-value for heterogeneity < 0.1), whereby the associations were positive only for distal colon cancer (***Supplementary Table 3***). In separate analyses by gender, heterogeneous associations were observed for total alcohol and spirits, for which the positive associations were only observed in men, and also for magnesium, fibre, and non-white bread for which the inverse associations were only observed in men (***Supplementary Table 4***). When we investigated the association of red and processed meat with CRC risk by follow-up duration, a trend towards smaller HRs was observed as follow-up increased (***Supplementary Figure 1***).

Replication analysis in NLCS

Of the 20 associations with an FDR<0.05 in EPIC, four associations reached nominal statistical significance in the NLCS cohort in the analysis for CRC (***Figure 2; <u>Supplementary Table 5</u>***), namely alcohol (HR = 1.06; 95%CI: 1.01-1.12), liquor/spirits (HR = 1.06; 95%CI: 1.01-1.11), milk (HR = 0.93; 95%CI: 0.89-0.98), and calcium intake (HR = 0.94; 95%CI: 0.90-0.99). An additional four associations, namely phosphorus, magnesium, riboflavin and total protein, were borderline significant in NLCS (HR for all four associations was: 0.95; 95%CI: 0.90-1.00) and the point estimates were almost identical to the ones calculated in EPIC.

In a separate analysis by tumour subsite in the NLCS, we found that most associations were consistent across the different subsites, with heterogeneous associations only evident for phosphorus (p-value for heterogeneity = 0.019), potassium (p = 0.014), vitamin B6 (p = 0.004), beta-carotene (p = 0.057) and total protein (p = 0.076) in the analysis for colon versus rectal cancer. The inverse associations of phosphorus, beta-carotene and total protein were only present for risk of colon cancer but not for rectal cancer. Associations for potassium and vitamin B6 were borderline statistically significantly inverse for colon cancer, but positive for rectal cancer (***<u>Supplementary Table 6</u>)***. Little heterogeneity was observed between proximal and distal colon cancer subsites (***<u>Supplementary Table 7</u>***), and by sex for CRC risk (***<u>Supplementary Table 8</u>).***

Meta-Analysis of EPIC and NLCS

The associations for most of the 20 dietary variables with CRC risk were homogeneous between EPIC and NLCS, except for soft drinks, vitamin B6, fruit, fibre, non-white bread, banana, and pork (p-value for heterogeneity <0.1 and/or I^2>50%), where the associations were null in NLCS and therefore a meta-analysis was not performed (***Figure 2; <u>Supplementary Table 5</u>)***. The remaining 13 associations yielded a nominally significant summary finding: alcohol (HR: 1.07, 95%CI: 1.04-1.09, I^2=0%), liquor/spirits (HR: 1.04, 95%CI: 1.02-1.06, I^2=0%), wine (HR: 1.04; 95%CI: 1.02-1.07; I^2 = 0%), beer/cider (HR: 1.06; 95%CI: 1.04-1.08; I^2 = 41%), milk (HR: 0.95, 95%CI: 0.93-0.98, I^2=26%), cheese (HR: 0.96; 95%CI: 0.94-0.99; I^2 = 33%), calcium (HR: 0.93, 95%CI: 0.90-0.95, I^2=0%), phosphorus (HR: 0.92; 95%CI: 0.90-0.95; I^2 = 29%), magnesium (HR: 0.95; 95%CI: 0.92-0.98; I^2 = 0%), potassium (HR: 0.96; 95%CI: 0.94-0.99; I^2 = 7%), riboflavin (HR: 0.94; 95%CI: 0.92-0.97; I^2 = 0%), beta-carotene (HR: 0.96; 95%CI: 0.93-0.98; I^2 = 0%), and total protein (HR: 0.94; 95%CI: 0.92-0.97; I^2 = 0%) (***Figure 2***).

Pairwise correlations and Mutual-adjustment analysis

The pair-wise correlation coefficients for the 20 FDR-significant foods/nutrients in EPIC ranged from -0.25 to 0.79 (***Supplementary Figure 2***). The largest coefficients (Spearman's $\rho > 0.50$) were: between alcohol and wine (ρ=0.79); between calcium and milk (ρ=0.53), phosphorus (ρ = 0.67), riboflavin (ρ = 0.64); between phosphorus and potassium (ρ=0.58), riboflavin (ρ=0.61), total protein (ρ=0.62); and between potassium and magnesium (ρ=0.61), riboflavin (ρ=0.54), vitamin B6 (ρ=0.66) and dietary fibre (ρ=0.51).

When alcohol, milk, cheese, calcium, phosphorus, magnesium, potassium, riboflavin, beta-carotene and total protein were included in a single multivariable-adjusted model in EPIC, only alcohol remained significantly associated with CRC risk (HR: 1.05; 95%CI: 1.03-1.11) (***Supplementary Table 9***)

Gene-Nutrient interaction analysis

Of the 73×20 gene-nutrient multiplicative interactions that were tested, considering a nominal p-value threshold (0.05), 85 were statistically significant in the analysis for CRC, 89 for colon cancer, 83 for rectal, 86 for proximal and 67 for distal colon cancer risk. Using the Bonferroni adjusted P-value threshold of 3.4×10^{-5}, no interaction remained significant (***Supplementary Table 10***).

Discussion

We used the NWAS approach to systematically evaluate the association between dietary intakes of 92 foods and nutrients and risk of CRC in EPIC and NLCS. We confirmed well-described associations in the literature for alcoholic beverages (positive), milk and calcium (inverse) with risk of CRC. In addition, our analysis showed that higher intakes of phosphorus, magnesium, potassium, riboflavin, beta-carotene, and total protein were associated with a lower risk of CRC.

Alcohol consumption was positively associated with risk of CRC in EPIC and NLCS, and this association was not different between colon and rectal cancer subsites or by type of alcoholic beverage. In agreement, the WCRF third Expert Report has graded the quality of this evidence as strong[21]. Persons with higher total alcohol consumption had a higher risk of CRC (summary HR per SD increment in daily intake: 1.07, 95% CI: 1.04-1.09), colon, and rectal cancer in the meta-analysis of EPIC and NLCS. When we evaluated this association by proximal vs. distal colon cancer and by sex, we found heterogeneous associations in EPIC, with associations only present for distal colon cancer and in men, but these findings were not confirmed in NLCS. The majority of the literature agrees that the positive association of alcohol consumption with CRC risk is consistent by anatomical subsite and sex[22-24]. Acetaldehyde, as a metabolite of ethanol oxidation, can be carcinogenic in colonocytes[25]. Higher ethanol consumption can induce oxidative stress, may act as a solvent for cellular penetration of other carcinogenic substances, can interfere with DNA repair mechanisms and negatively affects the gut flora synbiosis weakening the gut barrier function[21].

Our study also confirmed the inverse association between intake of **dairy products and calcium** with risk of CRC, where individuals with higher calcium consumption had a 7% lower risk of CRC per 334.5 mg increment in intake/day. One of the most prominent mechanisms by which calcium is thought to act to reduce CRC risk is by its ability to bind unconjugated bile acids and free fatty acids, diminishing their potential toxic effects on the colorectum[26]. Heterogeneity by anatomical subsite or gender was not observed, in agreement with the WCRF meta-analysis and a more recent publication in the Nurses' Health Study[21, 22]. Dairy products are also a rich source of **phosphorus**, which was also inversely associated with CRC risk in our study but has been infrequently studied in other publications. A previous analysis of nutrient patterns in EPIC identified a pattern characterised by total protein, riboflavin, phosphorus and calcium that was associated with a 4% decreased CRC risk[27]. All these nutrients were analysed independently in our analysis and yielded inverse associations in EPIC that were robust after correcting for multiple testing and were replicated in NLCS. Since several of these nutrients share common sources of intake, a correlation of approximately 0.50-0.70 was observed in EPIC, which makes it challenging to distinguish their independent effects[28].

Many studies have investigated the association between **red meat or processed meat** consumption and risk of CRC. A dose-response meta-analysis by the WCRF third Expert Report concluded that there is strong evidence that consuming red meat (including beef, pork, lamb and goat from domesticated animals) or processed meat (meat preserved by smoking, curing, salting or addition of chemical preservatives) increases the risk of CRC by 12 % per 100 g/d increment for red meat and 16% per 50 g/d for processed meat[4]. A combination of mechanisms may contribute to the higher risk of colorectal tumourigenesis among individuals consuming larger amounts of red and/or processed meat. Cooking meat at high temperatures may lead to the formation of heterocyclic amines (HCA) and polycyclic aromatic hydrocarbons (PAHs), which have been associated with colorectal carcinogenesis in experimental studies[29]. Red meat also contains haem iron at high levels that may stimulate the endogenous formation of carcinogenic N-nitroso compounds, which promote colorectal tumourigenesis[30]. Additionally, processed meat can be an exogenous source of N-nitroso compounds. Although accumulated evidence supports that higher intakes of red or processed meat are associated with higher risk of CRC, these findings were not replicated in our analysis in EPIC (HR per 36.2 grams of red meat intake daily: 1.02; 95%CI: 0.98-1.05; FDR: 0.507; HR per 31.5 grams of processed meat intake daily: 1.04; 95%CI: 1.00-1.08; FDR: 0.092). An earlier report from EPIC in 2005, with a mean follow-up of 4.8 years and 1,329 incident CRC cases, observed a positive association between red and processed meat consumption with CRC risk[31]. A potential reason for this discrepancy is that EPIC, as most other cohorts, has assessed meat consumption only during recruitment in the 1990s; thus, the current analysis assumes that consumption has stayed stable over two decades. However, a notable decrease in bovine meat consumption between 2000 and 2013 has been noticed in Europe[32], which was accompanied by an analogous increase in cheese, fish, dairy and poultry consumption. In the current paper, we observed a trend towards smaller HRs in the association of red and processed meat with CRC risk as follow-up increased. A recent time-varying exposure analysis in the Nurses' Health Study and the Health Professionals Follow-up Study showed that a decrease in red meat consumption and simultaneous increases in healthy alternative food choices over time were associated with a lower risk of all-cause mortality[33]. Additional reasons for the discrepant associations could be that stricter surveillance programmes and novel technologies have led to a relative decline in the nitrite content of meat products[34, 35].

The current NWAS study observed an inverse association of **magnesium** intake with risk of CRC, which agreed with the results of a recent meta-analysis of seven observational studies[36]. One purported mechanism by which magnesium may be implicated in lower CRC risk is by its potential to inhibit *c-myc* oncogene expression in colon cancer cells[37]. Furthermore, magnesium has been shown to improve insulin sensitivity and lower plasma insulin concentrations, which may have an impact on CRC development[38, 39].

We also observed an inverse association between intake of **beta-carotene** and risk of CRC, but few other studies have investigated this association[40, 41]. Our findings agree with a previous report from EPIC in 2014[41]. However, a cohort analysis in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) trial, comprising of 26,951 middle-aged male smokers, showed no association between dietary beta-carotene and risk of CRC[40].

Vitamins B2 and B6 are among the micronutrients that play a pivotal role in one-carbon metabolism, which has been related to carcinogenesis because of its involvement in the synthesis of purines and pyrimidines for subsequent DNA synthesis and in the synthesis of methionine for DNA methylation[42]. Additionally, deficiencies in these vitamins are common following high alcohol intake, which might act as an effect modifier in these associations. Inverse associations between **riboflavin** (vitamin B2) and **vitamin B6** intake and CRC risk were observed in EPIC, but only the association with riboflavin was replicated in the NLCS. Previous studies on the association between riboflavin intake and CRC risk are scarce[43]. Results from the Women's Health Initiative Observational Study indicated a 25% decreased CRC risk for the highest compared to the lowest quartile of total riboflavin intake, but was not statistically significant when only dietary intake of riboflavin was considered[43]. A meta-analysis of eight studies did not show an association between vitamin B6 intake and CRC risk, but blood levels of its active form, pyridoxal 5′-phosphate, were associated with lower CRC risk[44].

Little is known on the role that **potassium** may play in relation to CRC risk, and epidemiological evidence thus far is limited[45]. We cannot rule out the possibility that the inverse association observed in our study may mirror the effect of other nutrients, such as vitamin B6 or dietary fibre, which share common dietary sources with potassium.

We further investigated whether top hits from the NWAS analysis interact with top hits from GWAS for CRC, but we did not identify any robust interaction after adjusting for multiple comparisons. Similar null findings have been reported in previous investigations[46], but future studies with larger sample sizes, wider genome coverage and use of functional information to formulate relevant biochemical pathways are warranted[47].

Strengths of this study include its large size and long follow-up duration and the NWAS approach that involved a comprehensive assessment of foods and nutrients whilst accounting for multiplicity of tests and replication of findings in an independent cohort. Another strength was the ability to explore associations according to different anatomical subsites as well as by sex. The primary **limitation** was that the analysis relied on a single dietary assessment at recruitment, not allowing to capture potential changes in dietary habits over time. In addition, intercorrelations between dietary exposures and overall dietary patterns were not accounted for. Furthermore, it is possible that there might be an association for foods or nutrients that were not included in this analysis. Additionally, the discrepancies observed between EPIC and NLCS for some dietary exposures may be due to poor validation coefficients. However, among the exposures for which heterogeneity was observed, correlation between the baseline FFQs and 24-hour diet recalls was good for fruit, fibre, vitamin B6 and beverage consumption in NLCS and fairly good for fibre and fruit across most EPIC centres, and information was not available for non-white bread or vitamin B6 consumption[11, 48]. Finally, we cannot exclude the possibility of residual confounding, although we adjusted for several potential confounders.

In **conclusion**, our study confirmed the well-established positive association for alcohol consumption and inverse association for dairy products and calcium intake with CRC risk. The study further suggested that higher intakes of magnesium, phosphorus, potassium, riboflavin, beta-carotene and total protein are associated with lower CRC risk.

Funding

This work was supported by the World Cancer Research Fund International Regular Grant Programme (WCRF 2014/1180 to Konstantinos K. Tsilidis). EPIC is supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer for the coordination of EPIC. The national cohorts were supported by the Danish Cancer Society (Denmark); Ligue Contre le Cancer; Institut Gustave Roussy; Mutuelle Générale de l'Education Nationale and 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) and Statistics Netherlands (The Netherlands); European Research Council (ERC-2009-AdG 232997) and Nordforsk, 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 and Navarra and the Catalan Institute of Oncology (Barcelona) (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 and MR/M012190/1 to EPIC-Oxford) (United Kingdom). The funders had no role in the design and conduct of the study, the collection, analysis, and interpretation of the data, or the preparation, review, and approval of the manuscript, or in the decision to submit the manuscript for publication.

Conflict of interest: All authors declare no conflict of interest.

Disclaimer: Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Data availability: The EPIC study data can be accessed via an application to the EPIC Steering Committee (https://epic.iarc.fr/access/index.php). Further information is available from the corresponding author upon request.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians 2018;68:394-424.
2. Kyrgiou M, Kalliala I, Markozannes G, Gunter MJ, Paraskevaidis E, Gabra H, Martin-Hirsch P, Tsilidis KK. Adiposity and cancer at major anatomical sites: umbrella review of the literature. BMJ (Clinical research ed.) 2017;356:j477.
3. Rezende LFM, Sa TH, Markozannes G, Rey-Lopez JP, Lee IM, Tsilidis KK, Ioannidis JPA, Eluf-Neto J. Physical activity and cancer: an umbrella review of the literature including 22 major anatomical sites and 770 000 cancer cases. British journal of sports medicine 2018;52:826-33.
4. World Cancer Research Fund, American Institute for Cancer research. Food, Nutrition,Physical Activity, and the Prevention of Colorectal Cancer. Continuous Update Project Report 2018.
5. Patel CJ, Bhattacharya J, Butte AJ. An Environment-Wide Association Study (EWAS) on type 2 diabetes mellitus. Plos One 2010;5:e10746.
6. Papadimitriou N, Muller D, van den Brandt PA, Geybels M, Patel CJ, Gunter MJ, Lopez DS, Key TJ, Perez-Cornago A, Ferrari P, Vineis P, Weiderpass E, et al. A nutrient-wide association study for risk of prostate cancer in the European Prospective Investigation into Cancer and Nutrition and the Netherlands Cohort Study. European journal of nutrition 2019.
7. Heath AK, Muller DC, van den Brandt PA, Papadimitriou N, Critselis E, Gunter M, Vineis P, Weiderpass E, Fagherazzi G, Boeing H, Ferrari P, Olsen A, et al. Nutrient-wide association study of 92 foods and nutrients and breast cancer risk. Breast cancer research : BCR 2020;22:5.
8. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J, Overvad K, Tjonneland A, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr 2002;5:1113-24.
9. van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 1990;43:285-95.
10. Slimani N, Deharveng G, Unwin I, Southgate DA, Vignat J, Skeie G, Salvini S, Parpinel M, Moller A, Ireland J, Becker W, Farran A, et al. The EPIC nutrient database project (ENDB): a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. Eur J Clin Nutr 2007;61:1037-56.
11. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, Hermus RJ. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr 1994;48:253-65.
12. Goldbohm RA, van 't Veer P, van den Brandt PA, van 't Hof MA, Brants HA, Sturmans F, Hermus RJ. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. Eur J Clin Nutr 1995;49:420-9.
13. Nevo table: Dutch food composition table 1986-1987. The Hague, Netherlands. Voorlichtingsbureau voor de voeding 1986.
14. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol 1990;19:553-8.
15. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika 1986;73:1-11.
16. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. American journal of epidemiology 1986;124:17-27.

17. Ferrari P, Day NE, Boshuizen HC, Roddam A, Hoffmann K, Thiebaut A, Pera G, Overvad K, Lund E, Trichopoulou A, Tumino R, Gullberg B, et al. The evaluation of the diet/disease relation in the EPIC study: considerations for the calibration and the disease models. Int J Epidemiol 2008;37:368-78.

18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 1995;Ser B:289-300.

19. R Core Team. R: A language and environment for statistical computing, vol. 2016, 2018.

20. Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, Conti DV, Qu C, Jeon J, Edlund CK, Greenside P, Wainberg M, et al. Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet 2019;51:76-87.

21. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Diet, nutrition, physical activity and colorectal cancer. Available at dietandcancerreport.org.

22. Wei EK, Colditz GA, Giovannucci EL, Wu K, Glynn RJ, Fuchs CS, Stampfer M, Willett W, Ogino S, Rosner B. A Comprehensive Model of Colorectal Cancer by Risk Factor Status and Subsite Using Data From the Nurses' Health Study. American journal of epidemiology 2017;185:224-37.

23. Burón Pust A, Alison R, Blanks R, Pirie K, Gaitskell K, Barnes I, Gathani T, Reeves G, Beral V, Green J. Heterogeneity of colorectal cancer risk by tumour characteristics: Large prospective study of UK women. International journal of cancer 2017;140:1082-90.

24. Murphy N, Ward HA, Jenab M, Rothwell JA, Boutron-Ruault MC, Carbonnel F, Kvaskoff M, Kaaks R, Kühn T, Boeing H, Aleksandrova K, Weiderpass E, et al. Heterogeneity of Colorectal Cancer Risk Factors by Anatomical Subsite in 10 European Countries: A Multinational Cohort Study. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2019;17:1323-31.e6.

25. Seitz HK, Stickel F. Molecular mechanisms of alcohol-mediated carcinogenesis. Nature reviews. Cancer 2007;7:599-612.

26. Newmark HL, Wargovich MJ, Bruce WR. Colon cancer and dietary fat, phosphate, and calcium: a hypothesis. J Natl Cancer Inst 1984;72:1323-5.

27. Moskal A, Freisling H, Byrnes G, Assi N, Fahey MT, Jenab M, Ferrari P, Tjønneland A, Petersen KE, Dahm CC, Hansen CP, Affret A, et al. Main nutrient patterns and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition study. British journal of cancer 2016;115:1430-40.

28. Welch AA, Fransen H, Jenab M, Boutron-Ruault MC, Tumino R, Agnoli C, Ericson U, Johansson I, Ferrari P, Engeset D, Lund E, Lentjes M, et al. Variation in intakes of calcium, phosphorus, magnesium, iron and potassium in 10 countries in the European Prospective Investigation into Cancer and Nutrition study. Eur J Clin Nutr 2009;63 Suppl 4:S101-21.

29. Cross AJ, Sinha R. Meat-related mutagens/carcinogens in the etiology of colorectal cancer. Environmental and molecular mutagenesis 2004;44:44-55.

30. Santarelli RL, Pierre F, Corpet DE. Processed meat and colorectal cancer: a review of epidemiologic and experimental evidence. Nutrition and cancer 2008;60:131-44.

31. Norat T, Bingham S, Ferrari P, Slimani N, Jenab M, Mazuir M, Overvad K, Olsen A, Tjonneland A, Clavel F, Boutron-Ruault MC, Kesse E, et al. Meat, fish, and colorectal cancer risk: the European Prospective Investigation into cancer and nutrition. J Natl Cancer Inst 2005;97:906-16.

32. Database FaAOCS. Food Supply - Livestock and Fish Primary Equivalent.

33. Zheng Y, Li Y, Satija A, Pan A, Sotos-Prieto M, Rimm E, Willett WC, Hu FB. Association of changes in red meat consumption with total and cause specific mortality among US women and men: two prospective cohort studies. BMJ (Clinical research ed.) 2019;365:l2110.

34. Troy DJ, Ojha KS, Kerry JP, Tiwari BK. Sustainable and consumer-friendly emerging technologies for application within the meat industry: An overview. Meat Science 2016;120:2-9.

35. Authority EFS. Opinion of the Scientific Panel on biological hazards (BIOHAZ) related to the effects of Nitrites/Nitrates on the Microbiological Safety of Meat Products. EFSA Journal 2004;2:14.

36. Meng Y, Sun J, Yu J, Wang C, Su J. Dietary Intakes of Calcium, Iron, Magnesium, and Potassium Elements and the Risk of Colorectal Cancer: a Meta-Analysis 2019;189:325-35.

37. Mori H, Morishita Y, Shinoda T, Tanaka T. Preventive effect of magnesium hydroxide on carcinogen-induced large bowel carcinogenesis in rats. Basic life sciences 1993;61:111-8.

38. Paolisso G, Sgambato S, Pizza G, Passariello N, Varricchio M, D'Onofrio F. Improved insulin response and action by chronic magnesium administration in aged NIDDM subjects. Diabetes care 1989;12:265-9.

39. van den Brandt PA, Smits KM, Goldbohm RA, Weijenberg MP. Magnesium intake and colorectal cancer risk in the Netherlands Cohort Study. British journal of cancer 2007;96:510-3.

40. Malila N, Virtamo J, Virtanen M, Pietinen P, Albanes D, Teppo L. Dietary and serum alpha-tocopherol, beta-carotene and retinol, and risk for colorectal cancer in male smokers. Eur J Clin Nutr 2002;56:615-21.

41. Leenders M, Leufkens AM, Siersema PD, van Duijnhoven FJ, Vrieling A, Hulshof PJ, van Gils CH, Overvad K, Roswall N, Kyro C, Boutron-Ruault MC, Fagerhazzi G, et al. Plasma and dietary carotenoids and vitamins A, C and E and risk of colon and rectal cancer in the European Prospective Investigation into Cancer and Nutrition. International journal of cancer 2014;135:2930-9.

42. Newman AC, Maddocks ODK. One-carbon metabolism in cancer. British journal of cancer 2017;116:1499-504.

43. Zschäbitz S, Cheng T-YD, Neuhouser ML, Zheng Y, Ray RM, Miller JW, Song X, Maneval DR, Beresford SAA, Lane D, Shikany JM, Ulrich CM. B vitamin intakes and incidence of colorectal cancer: results from the Women's Health Initiative Observational Study cohort. The American journal of clinical nutrition 2013;97:332-43.

44. Larsson SC, Orsini N, Wolk A. Vitamin B6 and risk of colorectal cancer: a meta-analysis of prospective studies. Jama 2010;303:1077-83.

45. Key TJ, Appleby PN, Masset G, Brunner EJ, Cade JE, Greenwood DC, Stephen AM, Kuh D, Bhaniani A, Powell N, Khaw K-T. Vitamins, minerals, essential fatty acids and colorectal cancer risk in the United Kingdom Dietary Cohort Consortium. International journal of cancer 2012;131:E320-E25.

46. Yang T, Li X, Montazeri Z, Little J, Farrington SM, Ioannidis JPA, Dunlop MG, Campbell H, Timofeeva M, Theodoratou E. Gene-environment interactions and colorectal cancer risk: An umbrella review of systematic reviews and meta-analyses of observational studies. International journal of cancer 2019;145:2315-29.

47. Figueiredo JC, Hsu L, Hutter CM, Lin Y, Campbell PT, Baron JA, Berndt SI, Jiao S, Casey G, Fortini B, Chan AT, Cotterchio M, et al. Genome-wide diet-gene interaction analyses for risk of colorectal cancer. PLoS genetics 2014;10:e1004228.

48. Kaaks R, Slimani N, Riboli E. Pilot phase studies on the accuracy of dietary intake measurements in the EPIC project: overall evaluation of results. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol 1997;26 Suppl 1:S26-36.

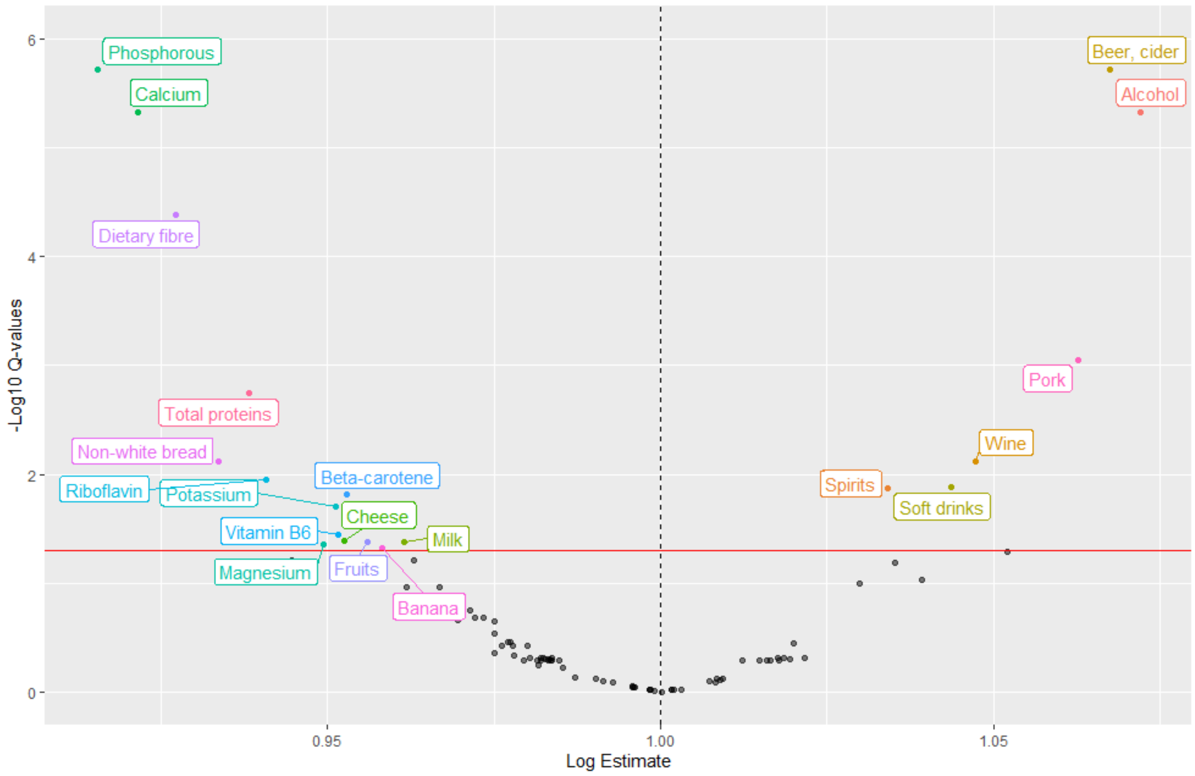


Figure 1. Volcano plot showing results from the nutrient-wide association study regarding the association between 92 dietary factors and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition study. The Y-axis shows the false discovery rate (FDR) adjusted P-values in –log10 scale from the Cox proportional hazards models for each dietary factor. The X-axis shows the estimated hazard ratio for each dietary factor per 1 standard deviation (SD) increase in daily consumption. The horizontal line represents the level of significance corresponding to FDR of 5%. The models were adjusted for total energy intake (kcal, continuous); smoking status (never, former, current); BMI (<20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥35kg/m^2); physical activity (inactive, moderately inactive, moderately active, active); diabetes history (no, yes); education status (none/primary, technical/professional, secondary, longer); and stratified by sex, age at recruitment (5-year intervals), and centre.

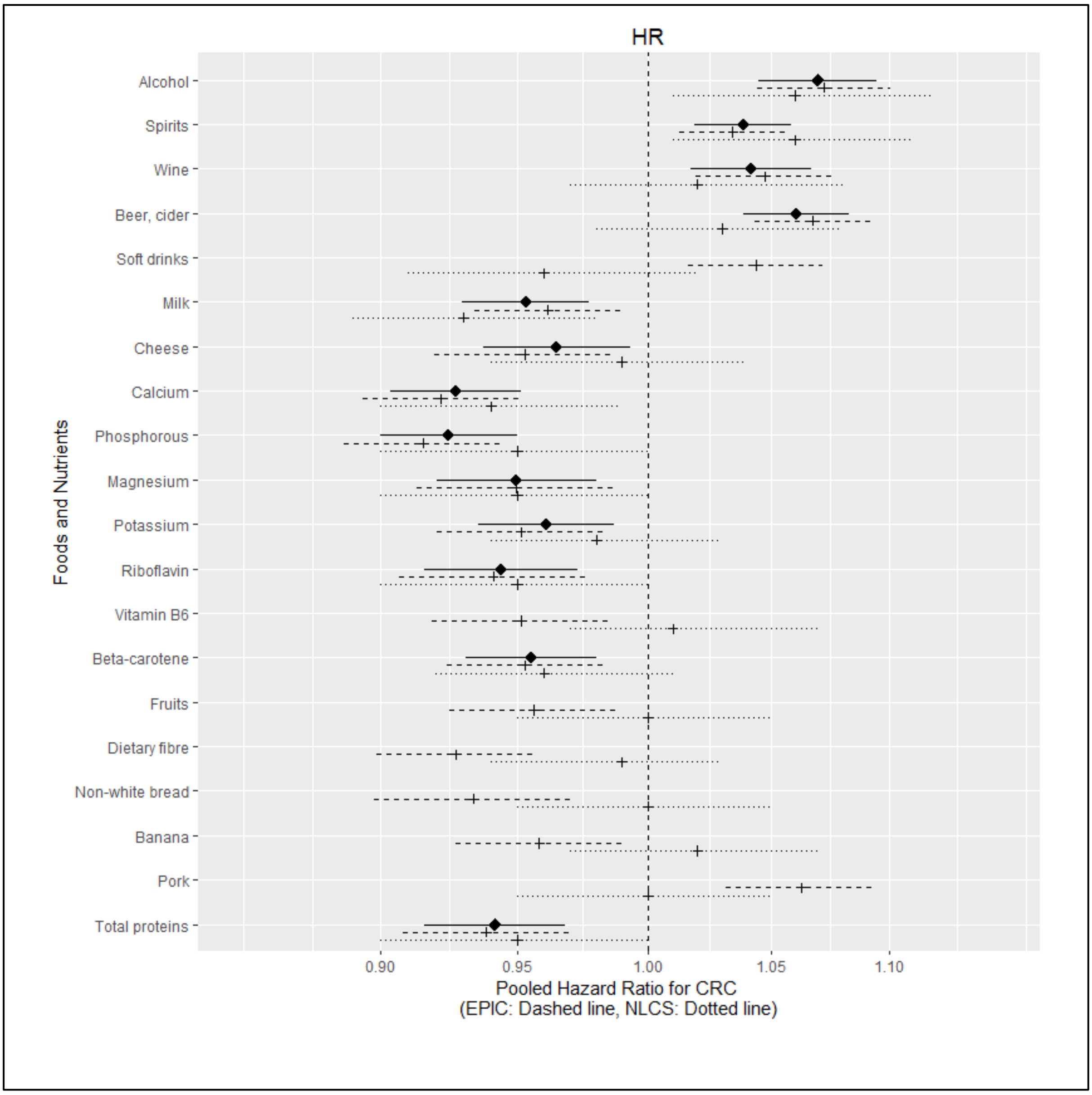


Figure 2. Forest plot showing the hazard ratios and 95% confidence intervals for the 20 FDR significant associations (FDR less than 5%), in the European Prospective Investigation into Cancer and Nutrition (EPIC) and the Netherlands Cohort Study (NLCS), as well as the results from a random effects meta-analysis on the two cohorts. The X-axis shows the estimated hazard ratio for each dietary factor for 1 standard deviation increase in daily consumption. The diamond and the solid line represent the pooled hazard ratio and 95%CI of the meta-analysis. The dashed (---) and the dotted (···) lines represent the results from the EPIC and the NLCS studies respectively. Meta-analysis was not performed when heterogeneity was high (p-value for heterogeneity < 0.1 and/or I^2 > 50%).

Table 1: Baseline demographic characteristics in EPIC and the NLCS subcohort.

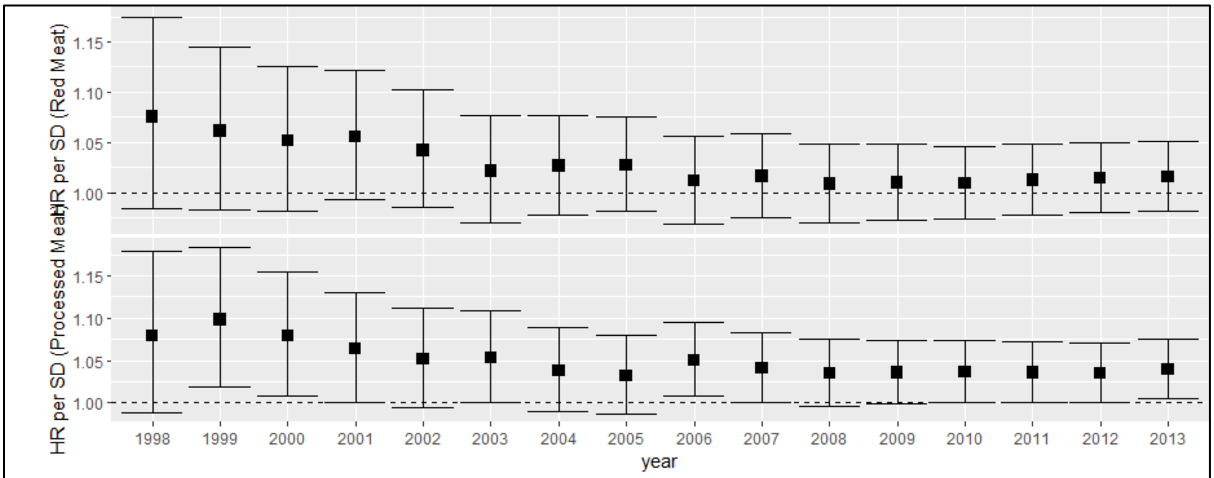
		EPIC			NLCS		
		Total, n (%)	**Non-cases, n (%)**	**Cases, n (%)**	**Total, n (%)**	**Non-cases, n (%)**	**Cases, n (%)**
	Total	386,792	381,723	5,069	7,496	3,731	3,765
Gender	**Male**	112,788 (29.2)	110,597 (29.0)	2,191 (43.2)	4,023 (53.7)	1,871 (50.1)	2,152 (57.2)
	Female	274,004 (70.8)	271,126 (71.0)	2,878 (56.8)	3,473 (46.3)	1,860 (49.9)	1,613 (42.8)
Age at recruitment (years)	**[<40)**	47,425 (12.3)	47,331 (12.4)	94 (1.9)	-	-	-
	[40, 45)	52,795 (13.6)	52,548 (13.8)	247 (4.9)	-	-	-
	[45, 50)	68,307 (17.7)	67,778 (17.8)	529 (10.4)	-	-	-
	[50, 55)	88,025 (22.8)	86,807 (22.7)	1,218 (24.0)	-	-	-
	[55, 60)	64,757 (16.7)	63,557 (16.7)	1,200 (23.7)	2,718 (36.3)	1,446 (38.8)	1,272 (33.8)
	[60, 65)	49,840 (12.9)	48,519 (12.7)	1,321 (26.1)	2,658 (35.5)	1,273 (34.1)	1,385 (36.8)
	[65, 70)	12,218 (3.2)	11,884 (3.1)	334 (6.6)	2,120 (28.3)	1,012 (27.1)	1,108 (29.4)
	[70, 75)	3,011 (0.8)	2,900 (0.8)	111 (2.2)			
	[>75]	414 (0.1)	399 (0.1)	15 (0.3)			
Smoking status	**Never**	194,087 (50.2)	191,990 (50.3)	2,097 (41.4)	2,474 (33.0)	1,303 (34.9)	1,171 (31.1)
	Former	103,942 (26.9)	102,268 (26.8)	1,674 (33)	2,991 (39.9)	1,364 (36.6)	1,627 (43.2)
	Current	88,763 (22.9)	87,465 (22.9)	1,298 (25.6)	2,031 (27.1)	1,064 (28.5)	967 (25.7)
Education[1]	**None/primary School**	112,507 (29.1)	110,607 (29.0)	1,900 (37.5)	2,040 (27.2)	1,038 (27.8)	1,002 (26.6)
	Technical/ professional school	87,563 (22.6)	86,290 (22.6)	1,273 (25.1)	1,599 (21.3)	798 (21.4)	801 (21.3)
	Secondary school	86,072 (22.3)	85,224 (22.3)	848 (16.7)	2,697 (36.0)	1,349 (36.2)	1,348 (35.8)
	Longer education (incl. university degree)	100,650 (26.0)	99,602 (26.1)	1,048 (20.7)	1,160 (15.5)	546 (14.6)	614 (16.3)
BMI (kg/m^2)	**[<20)**	26,550 (6.9)	26,385 (6.9)	165 (3.3)	243 (3.2)	139 (3.7)	104 (2.8)
	[20, 23)	99,036 (25.6)	98,100 (25.7)	936 (18.5)	1,528 (20.4)	783 (21.0)	745 (19.8)
	[23, 25)	81,112 (21.0)	80,111 (21.0)	1,001 (19.7)	2,231 (29.8)	1,129 (30.3)	1,102 (29.3)
	[25, 30)	131,871 (34.1)	129,747 (34.0)	2,124 (41.9)	3,037 (40.5)	1,445 (38.7)	1,592 (42.3)
	[30, 35)	38,125 (9.9)	37,464 (9.8)	661 (13.0)	403 (5.4)	208 (5.6)	195 (5.2)
	[>35]	10,098 (2.6)	9,916 (2.6)	182 (3.6)	54 (0.7)	27 (0.7)	27 (0.7)
Physical activity[2]	**Inactive**	72,301 (18.7)	71,167 (18.6)	1,134 (22.4)	1,546 (20.6)	765 (20.5)	781 (20.7)
	Moderately inactive	132,369 (34.2)	130,641 (34.2)	1,728 (34.1)	2,350 (31.4)	1,172 (31.4)	1,178 (31.3)

	Moderately active	106,613 (27.6)	105,417 (27.6)	1,196 (23.6)	1,623 (21.7)	798 (21.4)	825 (21.9)
	Active	75,509 (19.5)	74,498 (19.5)	1,011 (19.9)	1,977 (26.4)	996 (26.7)	981 (26.1)
Diabetes	**No**	376,678 (97.4)	371,832 (97.4)	4,846 (95.6)	7,271 (97.0)	3,608 (96.7)	3,663 (97.3)
	Yes	10,114 (2.6)	9,891 (2.6)	223 (4.4)	225 (3.0)	123 (3.3)	102 (2.7)
Family history of CRC	**No**	-	-	-	6,935 (92.5)	3,527 (94.5)	3,408 (90.5)
	Yes	-	-	-	561 (7.5)	204 (5.5)	357 (9.5)

EPIC: European Prospective Investigation into Cancer and Nutrition; NLCS: Netherlands Cohort Study; BMI: body mass index; CRC: colorectal cancer

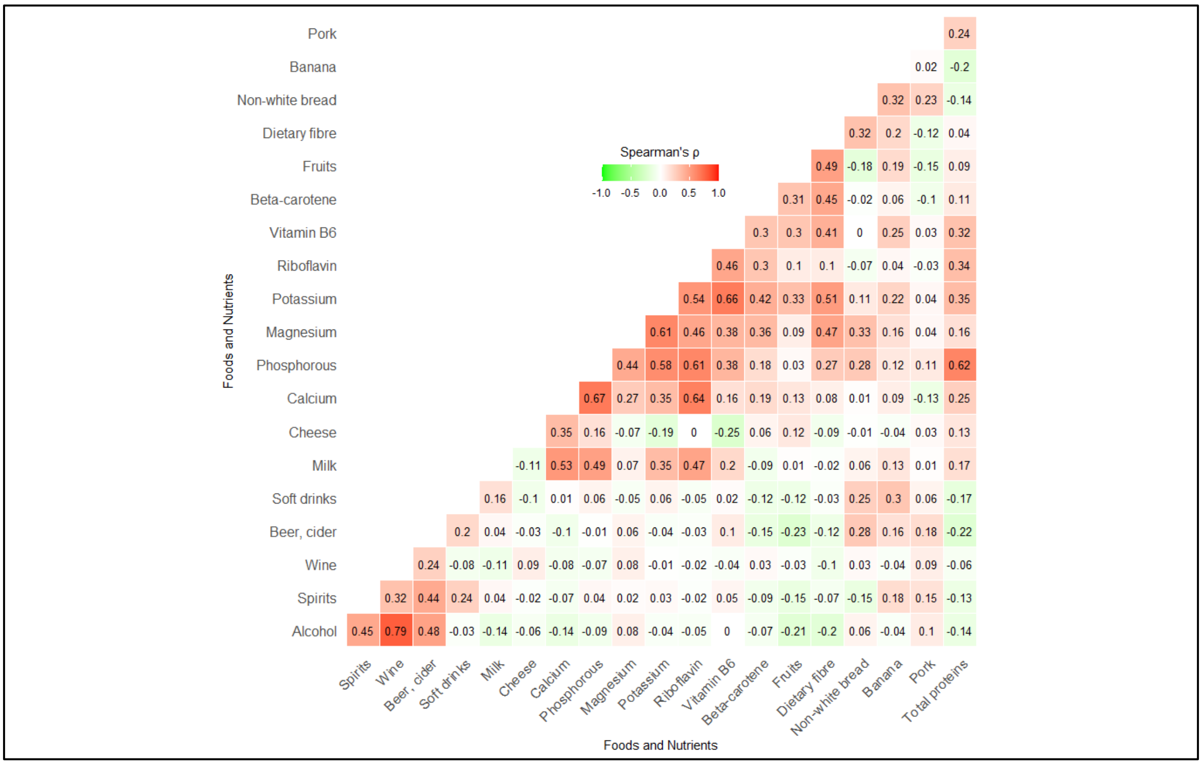
[1]The four educational level categories in NLCS were formed as follows: Primary school; Lower vocational school; Secondary, medium vocational; Higher vocational, university.

[2]The four physical activity categories in NLCS were based on non-occupational physical activity and formed as follows: <=30 min/d; >30-<=60 min/d; >60-<=90 min/d; >90 min/d.



	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
events	668	919	1179	1481	1806	2129	2460	2830	3182	3574	3927	4283	4603	4847	5031	5069
mean_py	3.26	4.23	5.18	6.13	7.05	7.97	8.87	9.74	10.59	11.4	12.13	12.78	13.34	13.72	14.05	14.09

***Supplementary Figure 1*.** Estimated hazard ratio of red meat (top panel) and processed meat (bottom panel) in relation to CRC risk in EPIC, per cumulative year of follow up. The Y-axis shows the estimated hazard ratio for each dietary factor for 1 standard deviation increase in daily consumption. The models were adjusted for total energy intake (kcal, continuous); smoking status (never, former, current); BMI (<20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥35kg/m^2); physical activity (inactive, moderately inactive, moderately active, active); diabetes history (no, yes); education status (none/primary, technical/professional, secondary, longer); and stratified by sex, age at recruitment (5-year intervals), and centre.



__Supplementary Figure 2__. *Pairwise* partial correlation coefficients (Spearman's ρ) of the 20 FDR-significant foods/nutrients in EPIC, adjusting for age, sex and centre.

Supplemental table 1. Hazard ratios[a] and 95% CIs for the association of **92 food and nutrient** intakes in relation to **colorectal cancer** risk in the EPIC study.

Dietary Variables	**HR (95%CI)[a]**	**P-value**	**FDR**	**SD**
Alcohol (g)	1.07 (1.04-1.10)	<0.001	<0.001	17.8
Spirits (g)[b]	1.03 (1.01-1.06)	0.002	0.013	12.2
Wine (g)	1.05 (1.02-1.08)	0.001	0.008	133.0
Beer, cider (g)	1.07 (1.04-1.09)	<0.001	<0.001	244.0
Soft drinks (g)	1.04 (1.02-1.07)	0.002	0.013	165.8
Milk (g)	0.96 (0.93-0.99)	0.008	0.041	208.3
Cheese (g)	0.95 (0.92-0.99)	0.007	0.041	34.2
Calcium (mg)	0.92 (0.89-0.95)	<0.001	<0.001	334.5
Phosphorous (mg)	0.92 (0.89-0.94)	<0.001	<0.001	273.6
Magnesium (mg)	0.95 (0.91-0.99)	0.009	0.044	82.3
Potassium (mg)	0.95 (0.92-0.98)	0.003	0.020	717.2
Riboflavin (mg)	0.94 (0.91-0.98)	0.001	0.011	0.6
Vitamin B6 (mg)	0.95 (0.92-0.99)	0.006	0.035	0.4
Beta-carotene (μg)	0.95 (0.92-0.98)	0.002	0.015	2,780.8
Fruits (g)	0.96 (0.92-0.99)	0.008	0.041	178.1
Dietary fibre (g)	0.93 (0.90-0.96)	<0.001	<0.001	6.2
Non-white bread (g)	0.93 (0.90-0.97)	0.001	0.008	72.9
Banana (g)	0.96 (0.93-0.99)	0.01	0.048	36.9
Pork (g)	1.06 (1.03-1.09)	<0.001	0.001	17.5
Total proteins (g)	0.94 (0.91-0.97)	<0.001	0.002	15.5
White bread (g)	1.05 (1.01-1.09)	0.012	0.052	73.6
Legumes (g)[c]	0.94 (0.90-0.99)	0.015	0.061	26.1
Root vegetables (g)	0.96 (0.93-0.99)	0.016	0.062	30.2
Protein (animal; g)	0.96 (0.93-0.99)	0.016	0.062	18.4
Eggs (g)[d]	1.04 (1.01-1.07)	0.018	0.064	17.5
Processed meat (g)	1.04 (1.00-1.08)	0.026	0.092	31.5
Ice cream (g)	1.03 (1.00-1.06)	0.029	0.100	11.3
Fish (g)	0.96 (0.93-1.00)	0.033	0.109	31.0
Fatty fish (g)[e]	0.97 (0.94-1.00)	0.034	0.109	14.5
Fish products (g)[f]	0.97 (0.93-1.00)	0.045	0.137	8.7
Protein (plant) (g)	0.96 (0.93-1.00)	0.055	0.163	7.8
Carbohydrates (g)	0.97 (0.94-1.00)	0.061	0.176	36.9
Breakfast cereals (g)[g]	0.97 (0.94-1.00)	0.065	0.181	42.8
Apple, pear (g)	0.97 (0.95-1.00)	0.076	0.205	85.4
Saturated fats (g)	0.97 (0.94-1.00)	0.078	0.206	7.7
Berries (g)[h]	0.97 (0.94-1.00)	0.085	0.216	12.5
Total fats (g)	0.98 (0.95-1.00)	0.090	0.223	13.5
Iron (mg)	0.98 (0.94-1.01)	0.120	0.291	2.6
Vitamin C (mg)	0.98 (0.95-1.01)	0.151	0.348	60.7
Total sugars (g)	0.98 (0.95-1.01)	0.151	0.348	32.3
Confectionery (non-chocolate; g)[i]	1.02 (0.99-1.05)	0.158	0.354	12.4

Monounsaturated fats (g)	0.98 (0.94-1.01)	0.177	0.371	7.3
Starch (g)	0.98 (0.95-1.01)	0.173	0.371	32.6
Yoghurt (g)	0.98 (0.95-1.01)	0.176	0.371	89.0
Stone fruits (g)[j]	0.98 (0.94-1.01)	0.214	0.437	45.6
Bread (g)	0.98 (0.94-1.01)	0.230	0.460	79.7
Mushrooms (g)[j]	1.02 (0.98-1.06)	0.267	0.482	9.0
Liver (g)[k]	1.02 (0.99-1.05)	0.259	0.482	4.7
Sugars (Sugar, honey, jam and syrup; g)	0.98 (0.95-1.01)	0.250	0.482	20.3
Fats (animal; g)	0.98 (0.95-1.01)	0.261	0.482	13.0
Nuts (g)	0.98 (0.95-1.02)	0.265	0.482	8.4
Cholesterol (mg)	1.02 (0.99-1.05)	0.280	0.486	115.8
Vitamin B12 (μg)	0.98 (0.95-1.01)	0.280	0.486	3.6
Cream puddings/ desserts (g)[l]	0.98 (0.95-1.01)	0.290	0.494	23.3
Soup (g)[m]	1.02 (0.98-1.06)	0.296	0.495	79.3
Citrus fruits (g)	0.98 (0.95-1.02)	0.302	0.497	62.5
Dry cakes, biscuits (g)[n]	0.98 (0.95-1.02)	0.313	0.506	12.3
Leafy vegetables (g)[j]	0.98 (0.94-1.02)	0.347	0.507	41.1
Mayonnaise (g)[o]	1.02 (0.98-1.06)	0.340	0.507	5.7
Fats (plant; g)	0.98 (0.94-1.02)	0.359	0.507	13.1
Red meat (g)	1.02 (0.98-1.05)	0.369	0.507	36.2
Crustaceans (g)[p]	1.01 (0.98-1.05)	0.347	0.507	6.1
Cakes, sweets (non-milk based; g)	0.98 (0.95-1.02)	0.344	0.507	38.5
Retinol (u)	1.01 (0.99-1.04)	0.365	0.507	694.9
Beef (g)[q]	0.98 (0.95-1.02)	0.340	0.507	19.2
Cabbage (g)[n]	0.98 (0.94-1.02)	0.367	0.507	30.9
Stalk vegetables, sprouts (g)[j]	1.02 (0.98-1.05)	0.355	0.507	12.8
Vitamin E (mg)	0.98 (0.95-1.02)	0.375	0.508	4.4
Thiamin (mg)	0.98 (0.94-1.03)	0.421	0.561	0.4
Lean fish (g)[r]	0.99 (0.95-1.02)	0.451	0.592	23.4
Grain and pod vegetables (g)[j]	0.99 (0.94-1.03)	0.563	0.730	12.7
Grapes (g)[s]	1.01 (0.98-1.04)	0.590	0.753	15.3
Potatoes (g)	1.01 (0.98-1.04)	0.607	0.755	74.8
Onion, garlic (g)[t]	0.99 (0.95-1.03)	0.605	0.755	14.7
Margarine (g)	1.01 (0.97-1.05)	0.621	0.762	16.2
Vitamin D (μg)	0.99 (0.96-1.03)	0.645	0.780	3.5
Offal (g)[u]	1.01 (0.97-1.04)	0.664	0.793	6.2
Crispbread, rusks (g)	0.99 (0.96-1.03)	0.688	0.801	17.1
Pasta, rice, other grains (g)	1.01 (0.97-1.05)	0.679	0.801	65.5
Fortified wines (g)[v]	1.00 (0.97-1.02)	0.768	0.884	15.8
Tea (g)[u]	1.00 (0.96-1.03)	0.809	0.907	304.1
Chocolate (g)	1.00 (0.96-1.03)	0.818	0.907	13.6
Coffee (g)	1.00 (0.96-1.03)	0.800	0.907	375.7
Sauces (g)[w]	1.00 (0.97-1.04)	0.865	0.948	18.9
Polyunsaturated fats (g)	1.00 (0.97-1.03)	0.924	0.955	4.5

Fruiting vegetables (g)[u]	1.00 (0.97-1.04)	0.923	0.955	52.8
Butter (g)	1.00 (0.97-1.03)	0.920	0.955	8.6
Margarine (vegetables; g)	1.00 (0.97-1.03)	0.899	0.955	13.1
Salty biscuits, crackers (g)	1.00 (0.97-1.03)	0.919	0.955	6.4
Fruit and vegetables juice (g)	1.00 (0.96-1.03)	0.935	0.956	115.3
Poultry (g)	1.00 (0.97-1.03)	0.952	0.962	19.8
Lamb (g)[x]	1.00 (0.97-1.04)	0.991	0.991	7.9

[a]All dietary factors entered the models as standardized continuous variables and reflect associations **per one standard deviation increase in daily consumption**. Nutrient intakes were adjusted for total energy intake using the regression residual method. The models were adjusted for total energy intake (kcal, continuous); smoking status (never, former, current); BMI <20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥35kg/m2); physical activity (inactive, moderately inactive, moderately active, active); diabetes history (no, yes); education status (none/primary, technical/professional, secondary, longer [including university]). They were further stratified by age at recruitment (<40, 40-44.9, 45-49.9, 50-54.9, 55-59.9, 60-64.9, 65-69.9, 70-74.9, ≥75), sex and recruitment centre.

[b]Intake of spirits was missing for participants from Italy and Norway (9.2 % missing across EPIC).
[c]Intake of legumes was missing for participants from Denmark and Norway (20.0% missing across EPIC).
[d]Intake of egg was missing for participants from Sweden (6.1% missing across EPIC).
[e]Intake of fatty fish was missing for participants from Germany (6.6% missing across EPIC).
[f]Intake of fish products was missing for participants from France and Italy (24.4% missing across EPIC).
[g]Intake of breakfast cereals was missing for participants from Italy (10.2% missing across EPIC).
[h]Intake of berries was missing for participants from Norway and the United Kingdom (16.6% missing across EPIC).
[i]Intake of confectionary was missing for participants from Germany and Norway (19.0% missing across EPIC).
[j]Intake for mushrooms, leafy vegetables, stone fruits, stalk vegetables, pod vegetables was missing for participants from Norway and Sweden (12.6% missing across EPIC).
[k]Intake of liver was missing for participants from The Netherlands, Norway and Sweden (20.7% missing across EPIC).
[l]Intake of cream puddings/desserts was missing for participants from Italy and Sweden (17.6% missing across EPIC).
[m]Intake of soup was missing for participants from Denmark, Italy and Norway (21.2% missing across EPIC).
[n]Intake of cabbage and biscuits was missing for participants from Sweden (6.1% missing across EPIC).
[o]Intake of mayonnaise was missing for participants from Italy, Norway and Sweden (13.9% missing across EPIC).
[p]Intake of crustaceans was missing for participants from Germany (12.5% missing across EPIC).
[q]Intake of beef was missing for participants from Sweden (6.1% missing across EPIC).
[r]Intake of lean fish was missing for participants from Germany, Italy and Sweden (19.9% missing across EPIC).
[s]Intake of grapes was missing for participants from Norway and Sweden (26.1% missing across EPIC).
[t]Intake for onion and garlic was missing for participants from France, Norway and Sweden (28.4% missing across EPIC).
[u]Intake of offal, tea and fruiting vegetables was missing for participants from Norway (6.4% missing across EPIC).
[v]Intake of fortified wines was missing for participants from Italy, Norway and Sweden (15.4% missing across EPIC).
[w]Intake of sauces was missing for participants from Italy (1.3% missing across EPIC).
[x]Intake of lamb was missing for participants from The Netherlands, Italy and Sweden (22.9% missing across EPIC).

Supplemental table 2. Hazard ratios[a] and 95% CIs for the association of the 20 food and nutrient intakes with colorectal cancer risk by **tumour location (colon vs rectal)** in the EPIC study.

Dietary Variables	**Colon, HR (95%CI)[a]**	**Rectum, HR (95%CI)[a]**	**P-value for Heterogeneity**
Alcohol	1.06 (1.02-1.10)	1.09 (1.04-1.14)	0.309
Spirits[b]	1.02 (0.99-1.05)	1.06 (1.02-1.09)	0.113
Wine	1.04 (1.00-1.07)	1.07 (1.02-1.12)	0.278
Beer, cider	1.06 (1.02-1.09)	1.07 (1.04-1.11)	0.509
Soft drinks	1.04 (1.01-1.08)	1.04 (1.00-1.09)	0.988
Milk	0.96 (0.92-0.99)	0.96 (0.91-1.01)	0.946
Cheese	0.96 (0.92-1.01)	0.94 (0.88-1.00)	0.491
Calcium	0.93 (0.89-0.96)	0.91 (0.86-0.96)	0.552
Phosphorous	0.91 (0.87-0.95)	0.92 (0.87-0.97)	0.790
Magnesium	0.91 (0.87-0.96)	1.01 (0.95-1.08)	0.011
Potassium	0.92 (0.88-0.96)	1.01 (0.95-1.07)	0.008
Riboflavin	0.92 (0.88-0.97)	0.96 (0.90-1.02)	0.344
Vitamin B6	0.91 (0.87-0.95)	1.02 (0.96-1.08)	0.002
Beta-carotene	0.95 (0.91-0.98)	0.96 (0.91-1.01)	0.602
Fruits	0.95 (0.91-0.99)	0.97 (0.92-1.03)	0.569
Dietary fibre	0.92 (0.88-0.95)	0.95 (0.90-1.00)	0.306
Non-white bread	0.93 (0.88-0.98)	0.95 (0.89-1.01)	0.669
Banana	0.94 (0.90-0.98)	1.00 (0.95-1.06)	0.041
Pork	1.06 (1.01-1.10)	1.07 (1.02-1.12)	0.686
Total proteins	0.93 (0.89-0.97)	0.95 (0.90-1.01)	0.409

[a]All dietary factors entered the models as standardized continuous variables and reflect associations per one standard deviation increase in daily consumption. Nutrient intakes were adjusted for total energy intake using the regression residual method. The models were adjusted for total energy intake (kcal, continuous); smoking status (never, former, current); BMI <20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥35kg/m2); physical activity (inactive, moderately inactive, moderately active, active); diabetes history (no, yes); education status (none/primary, technical/professional, secondary, longer [including university]). They were further stratified by age at recruitment (<40, 40-44.9, 45-49.9, 50-54.9, 55-59.9, 60-64.9, 65-69.9, 70-74.9, ≥75), sex and recruitment centre.

[b]Intake of spirits was missing for participants from Italy and Norway (9.2% missing across EPIC).

Supplemental table 3. Hazard ratios[a] and 95% CIs for the association of the 20 food and nutrient intakes with colorectal cancer risk by **tumour location (proximal vs distal)** in the EPIC study.

Dietary Variables	**Proximal, HR (95%CI)[a]**	**Distal, HR (95%CI)[a]**	**P-value for Heterogeneity**
Alcohol	1.01 (0.96-1.07)	1.11 (1.05-1.16)	0.015
Spirits[b]	1.02 (0.98-1.07)	1.00 (0.96-1.05)	0.564
Wine	1.00 (0.95-1.06)	1.07 (1.02-1.12)	0.087
Beer, cider	1.04 (0.99-1.09)	1.08 (1.03-1.12)	0.298
Soft drinks	1.02 (0.97-1.08)	1.06 (1.01-1.11)	0.311
Milk	0.97 (0.92-1.02)	0.96 (0.91-1.02)	0.931
Cheese	0.98 (0.92-1.05)	0.93 (0.87-0.99)	0.245
Calcium	0.94 (0.89-0.99)	0.91 (0.86-0.97)	0.432
Phosphorous	0.93 (0.87-0.98)	0.90 (0.85-0.96)	0.546
Magnesium	0.96 (0.89-1.03)	0.88 (0.82-0.95)	0.138
Potassium	0.94 (0.88-1.00)	0.92 (0.86-0.98)	0.599
Riboflavin	0.95 (0.89-1.02)	0.90 (0.84-0.97)	0.309
Vitamin B6	0.90 (0.85-0.96)	0.94 (0.88-1.01)	0.366
Beta-carotene	0.94 (0.88-0.99)	0.97 (0.91-1.02)	0.431
Fruits	0.95 (0.89-1.01)	0.97 (0.91-1.03)	0.659
Dietary fibre	0.94 (0.88-0.99)	0.91 (0.86-0.96)	0.457
Non-white bread	0.95 (0.88-1.02)	0.93 (0.86-1.00)	0.643
Banana	0.91 (0.86-0.97)	0.98 (0.92-1.04)	0.128
Pork	1.03 (0.97-1.09)	1.08 (1.02-1.14)	0.263
Total proteins	0.92 (0.86-0.98)	0.93 (0.88-0.99)	0.702

[a]All dietary factors entered the models as standardized continuous variables and reflect associations per one standard deviation increase in daily consumption. Nutrient intakes were adjusted for total energy intake using the regression residual method. The models were adjusted for total energy intake (kcal, continuous); smoking status (never, former, current); BMI <20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥35kg/m2); physical activity (inactive, moderately inactive, moderately active, active); diabetes history (no, yes); education status (none/primary, technical/professional, secondary, longer [including university]). They were further stratified by age at recruitment (<40, 40-44.9, 45-49.9, 50-54.9, 55-59.9, 60-64.9, 65-69.9, 70-74.9, ≥75), sex and recruitment centre.

[b]Intake of spirits was missing for participants from Italy and Norway (9.2% missing across EPIC).

Supplemental table 4. Hazard ratios[a] and 95% CIs for the association of the 20 food and nutrient intakes with colorectal cancer risk by **sex (men vs women)** in the EPIC study.

Dietary Variables	**Men, HR (95%CI)[a]**	**Women, HR (95%CI)[a]**	**P-value for Heterogeneity**
Alcohol	1.12 (1.08-1.16)	1.03 (0.99-1.07)	0.002
Spirits[b]	1.05 (1.03-1.07)	0.98 (0.93-1.03)	0.010
Wine	1.04 (1.00-1.07)	1.06 (1.02-1.12)	0.386
Beer, cider	1.07 (1.05-1.10)	1.01 (0.93-1.10)	0.220
Soft drinks	1.03 (0.99-1.07)	1.06 (1.02-1.10)	0.376
Milk	0.96 (0.92-1.00)	0.97 (0.93-1.01)	0.777
Cheese	0.95 (0.90-1.00)	0.95 (0.91-1.00)	0.866
Calcium	0.91 (0.86-0.95)	0.93 (0.90-0.97)	0.407
Phosphorous	0.91 (0.86-0.95)	0.92 (0.89-0.96)	0.621
Magnesium	0.89 (0.84-0.96)	0.98 (0.93-1.03)	0.033
Potassium	0.92 (0.88-0.98)	0.97 (0.93-1.01)	0.170
Riboflavin	0.95 (0.89-1.01)	0.94 (0.90-0.98)	0.789
Vitamin B6	0.97 (0.92-1.02)	0.94 (0.90-0.98)	0.404
Beta-carotene	0.94 (0.88-0.99)	0.96 (0.93-1.00)	0.434
Fruits	0.95 (0.90-1.00)	0.96 (0.92-1.00)	0.688
Dietary fibre	0.88 (0.84-0.93)	0.96 (0.93-1.01)	0.006
Non-white bread	0.89 (0.84-0.94)	0.99 (0.94-1.05)	0.008
Banana	0.97 (0.92-1.01)	0.95 (0.91-0.99)	0.653
Pork	1.05 (1.01-1.09)	1.08 (1.03-1.14)	0.303
Total proteins	0.94 (0.89-0.99)	0.94 (0.90-0.98)	0.909

[a]All dietary factors entered the models as standardized continuous variables and reflect associations per one standard deviation increase in daily consumption. Nutrient intakes were adjusted for total energy intake using the regression residual method. The models were adjusted for total energy intake (kcal, continuous); smoking status (never, former, current); BMI <20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥35kg/m2); physical activity (inactive, moderately inactive, moderately active, active); diabetes history (no, yes); education status (none/primary, technical/professional, secondary, longer [including university]). They were further stratified by age at recruitment (<40, 40-44.9, 45-49.9, 50-54.9, 55-59.9, 60-64.9, 65-69.9, 70-74.9, ≥75), sex and recruitment centre.

[b]Intake of spirits was missing for participants from Italy and Norway (9.2% missing across EPIC).

Supplemental table 5. Hazard ratios and 95% CIs for the association of the 20 food and nutrient intakes with colorectal cancer risk in the EPIC and the NLCS study.

Dietary Variables	**EPIC study, HR[a] (95%CI)**	**NLCS study, HR[b] (95%CI)**	**P-value for heterogeneity**
Alcohol	1.07 (1.04-1.10)	1.06 (1.01-1.12)	0.704
Spirits[c]	1.03 (1.01-1.06)	1.06 (1.01-1.11)	0.350
Wine	1.05 (1.02-1.08)	1.02 (0.97-1.08)	0.389
Beer, cider	1.07 (1.04-1.09)	1.03 (0.98-1.08)	0.192
Soft drinks	1.04 (1.02-1.07)	0.96 (0.91-1.02)	0.009
Milk	0.96 (0.93-0.99)	0.93 (0.89-0.98)	0.245
Cheese	0.95 (0.92-0.99)	0.99 (0.94-1.04)	0.221
Calcium	0.92 (0.89-0.95)	0.94 (0.90-0.99)	0.494
Phosphorus	0.92 (0.89-0.94)	0.95 (0.90-1.00)	0.237
Magnesium	0.95 (0.91-0.99)	0.95 (0.90-1.00)	0.986
Potassium	0.95 (0.92-0.98)	0.98 (0.94-1.03)	0.300
Riboflavin	0.94 (0.91-0.98)	0.95 (0.90-1.00)	0.768
Vitamin B6	0.95 (0.92-0.99)	1.01 (0.97-1.07)	0.053
beta-carotene	0.95 (0.92-0.98)	0.96 (0.92-1.01)	0.795
Fruit	0.96 (0.92-0.99)	1.00 (0.95-1.05)	0.142
Fibre	0.93 (0.90-0.96)	0.99 (0.94-1.03)	0.021
Non-white bread	0.93 (0.90-0.97)	1.00 (0.95-1.05)	0.035
Bananas	0.96 (0.93-0.99)	1.02 (0.97-1.07)	0.038
Pork	1.06 (1.03-1.09)	1.00 (0.95-1.05)	0.040
Total protein	0.94 (0.91-0.97)	0.95 (0.90-1.00)	0.692

[a]All dietary factors entered the models as standardized continuous variables and reflect associations per one standard deviation increase in daily consumption. Nutrient intakes were adjusted for total energy intake using the regression residual method. The models were adjusted for total energy intake (kcal, continuous); smoking status (never, former, current); BMI <20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥35kg/m2); physical activity (inactive, moderately inactive, moderately active, active); diabetes history (no, yes); education status (none/primary, technical/professional, secondary, longer [including university]). They were further stratified by age at recruitment (<40, 40-44.9, 45-49.9, 50-54.9, 55-59.9, 60-64.9, 65-69.9, 70-74.9, ≥75), sex and recruitment centre.

[b]Multivariable analyses were stratified for age at baseline (55-59, 60-64, 65-69 yrs), sex, and adjusted for: smoking status (never, ex, current), BMI (<20, 20-<23, 23-<25, 25-<30, 30-<35, ≥35 kg/m2), non-occupational physical activity (≤30, >30-60, >60-90, >90 min/day), highest level of education (primary school or lower vocational, secondary or medium vocational, and higher vocational or university), family history of colorectal cancer (no, yes), history of diabetes, energy intake (kcal, continuous).

[c]Intake of spirits was missing for participants from Italy and Norway (9.2% missing across EPIC).

Supplemental table 6. Hazard ratios[a] and 95% CIs for the association of the 20 food and nutrients with colorectal cancer risk by **tumour location (colon vs rectal)** in the NLCS study.

Dietary Variables	**Colon, HR (95%CI)[a]**	**Rectum, HR (95%CI)[a]**	**P-value for heterogeneity**
Alcohol	1.03 (0.98-1.09)	1.11 (1.04-1.20)	0.100
Spirits	1.05 (0.99-1.10)	1.08 (1.00-1.16)	0.544
Wine	1.01 (0.95-1.06)	1.05 (0.97-1.14)	0.435
Beer, cider	0.99 (0.94-1.05)	1.06 (1.00-1.14)	0.118
Soft drinks	0.97 (0.92-1.03)	0.96 (0.87-1.06)	0.858
Milk	0.94 (0.89-0.99)	0.95 (0.88-1.03)	0.827
Cheese	0.98 (0.93-1.04)	1.04 (0.96-1.13)	0.239
Calcium	0.95 (0.90-1.00)	0.99 (0.91-1.07)	0.403
Phosphorus	0.93 (0.89-0.99)	1.04 (0.96-1.12)	0.019
Magnesium	0.94 (0.89-0.99)	1.00 (0.93-1.08)	0.186
Potassium	0.96 (0.91-1.01)	1.08 (1.00-1.17)	0.014
Riboflavin	0.94 (0.89-1.00)	1.00 (0.92-1.08)	0.221
Vitamin B6	0.98 (0.93-1.03)	1.12 (1.04-1.21)	0.004
beta-carotene	0.93 (0.88-0.98)	1.02 (0.94-1.10)	0.057
Fruits	0.99 (0.94-1.05)	1.02 (0.94-1.10)	0.543
Fibre	0.97 (0.92-1.03)	1.02 (0.95-1.10)	0.287
Non-white bread	0.98 (0.93-1.04)	1.06 (0.98-1.14)	0.102
Bananas	1.02 (0.96-1.08)	1.04 (0.96-1.12)	0.695
Pork	0.98 (0.93-1.03)	1.03 (0.95-1.12)	0.314
Total protein	0.93 (0.88-0.99)	1.02 (0.94-1.11)	0.076

[a]Multivariable analyses were stratified for age at baseline (55-59, 60-64, 65-69 yrs), sex, and adjusted for: smoking status (never, ex, current), BMI (<20, 20-<23, 23-<25, 25-<30, 30-<35, ≥35 kg/m2), non-occupational physical activity (≤30, >30-60, >60-90, >90 min/day), highest level of education (primary school or lower vocational, secondary or medium vocational, and higher vocational or university), family history of colorectal cancer (no, yes), history of diabetes, energy intake (kcal, continuous).

Supplemental table 7. Hazard ratios[a] and 95% CIs for the association of the 20 food and nutrient intakes with colorectal cancer risk by **tumour location (proximal vs distal)** in the NLCS study.

Dietary Variables	**Proximal, HR (95%CI)[a]**	**Distal, HR (95%CI)[a]**	**P-value for heterogeneity**
Alcohol	1.04 (0.97-1.11)	1.03 (0.96-1.10)	0.843
Spirits	1.05 (0.99-1.12)	1.04 (0.97-1.11)	0.837
Wine	1.01 (0.94-1.08)	1.00 (0.93-1.07)	0.843
Beer, cider	0.99 (0.92-1.06)	1.00 (0.93-1.07)	0.843
Soft drinks	0.97 (0.90-1.05)	0.98 (0.90-1.06)	0.858
Milk	0.94 (0.87-1.00)	0.95 (0.89-1.02)	0.831
Cheese	1.00 (0.93-1.07)	0.98 (0.91-1.06)	0.702
Calcium	0.96 (0.90-1.03)	0.95 (0.89-1.02)	0.831
Phosphorus	0.95 (0.89-1.01)	0.94 (0.87-1.00)	0.825
Magnesium	0.95 (0.89-1.02)	0.93 (0.87-1.00)	0.669
Potassium	0.97 (0.90-1.04)	0.96 (0.90-1.03)	0.837
Riboflavin	0.94 (0.88-1.01)	0.96 (0.90-1.03)	0.669
Vitamin B6	0.99 (0.92-1.06)	0.97 (0.91-1.05)	0.691
beta-carotene	0.96 (0.89-1.02)	0.89 (0.82-0.96)	0.154
Fruits	1.01 (0.94-1.08)	0.98 (0.91-1.05)	0.553
Fibre	0.98 (0.92-1.05)	0.97 (0.91-1.04)	0.831
Non-white bread	0.97 (0.91-1.05)	1.00 (0.93-1.07)	0.551
Bananas	1.01 (0.94-1.08)	1.03 (0.95-1.11)	0.712
Pork	0.97 (0.91-1.04)	0.99 (0.92-1.06)	0.681
Total protein	0.93 (0.86-1.00)	0.94 (0.87-1.01)	0.843

[a]Multivariable analyses were stratified for age at baseline (55-59, 60-64, 65-69 yrs), sex, and adjusted for: smoking status (never, ex, current), BMI (<20, 20-<23, 23-<25, 25-<30, 30-<35, ≥35 kg/m2), non-occupational physical activity (≤30, >30-60, >60-90, >90 min/day), highest level of education (primary school or lower vocational, secondary or medium vocational, and higher vocational or university), family history of colorectal cancer (no, yes), history of diabetes, energy intake (kcal, continuous).

Supplemental table 8. Hazard ratios[a] and 95% CIs for the association of the 20 food and nutrient intakes with colorectal cancer risk by **sex (men vs women)** in the NLCS study.

Dietary Variables	**Men, HR (95%CI)**	**Women, HR (95%CI)**	**P-value for heterogeneity**
Alcohol	1.07 (1.01-1.13)	1.06 (0.94-1.18)	0.848
Spirits	1.06 (1.01-1.12)	1.05 (0.90-1.23)	0.839
Wine	1.02 (0.95-1.09)	1.04 (0.96-1.12)	0.700
Beer, cider	1.02 (0.97-1.08)	0.99 (0.76-1.29)	0.557
Soft drinks	0.99 (0.91-1.07)	0.91 (0.83-1.01)	0.142
Milk	0.92 (0.86-0.98)	0.95 (0.88-1.03)	0.519
Cheese	1.01 (0.95-1.08)	0.95 (0.88-1.03)	0.247
Calcium	0.94 (0.88-1.01)	0.96 (0.89-1.03)	0.667
Phosphorus	0.96 (0.90-1.02)	0.94 (0.87-1.02)	0.661
Magnesium	0.95 (0.90-1.02)	0.95 (0.87-1.03)	1.000
Potassium	0.99 (0.92-1.05)	0.99 (0.92-1.07)	1.000
Riboflavin	0.94 (0.88-1.00)	0.97 (0.90-1.05)	0.523
Vitamin B6	1.02 (0.96-1.08)	1.02 (0.94-1.10)	1.000
beta-carotene	0.96 (0.90-1.02)	0.97 (0.90-1.04)	0.845
Fruits	1.01 (0.94-1.08)	0.99 (0.92-1.06)	0.694
Fibre	0.99 (0.94-1.05)	0.98 (0.91-1.06)	0.832
Non-white bread	1.00 (0.95-1.07)	1.00 (0.91-1.11)	1.000
Bananas	1.01 (0.94-1.08)	1.03 (0.96-1.11)	0.712
Pork	1.00 (0.94-1.07)	0.99 (0.92-1.07)	0.840
Total protein	0.96 (0.89-1.03)	0.94 (0.86-1.02)	0.697

[a]Multivariable analyses were stratified for age at baseline (55-59, 60-64, 65-69 yrs), sex, and adjusted for: smoking status (never, ex, current), BMI (<20, 20-<23, 23-<25, 25-<30, 30-<35, ≥35 kg/m2), non-occupational physical activity (≤30, >30-60, >60-90, >90 min/day), highest level of education (primary school or lower vocational, secondary or medium vocational, and higher vocational or university), family history of colorectal cancer (no, yes), history of diabetes, energy intake (kcal, continuous).

Supplemental table 9. Multivariable analysis of **mutually adjusted foods and nutrients.**

Variable	Beta*	SE	HR	Z-value	P-value	VIF
Alcohol	**0.0530**	**0.0140**	**1.0544**	**3.7784**	**0.0002**	1.2
Milk	0.0052	0.0244	1.0052	0.2133	0.8311	2.9
Cheese	-0.0013	0.0260	0.9987	-0.0493	0.9607	2.4
Calcium	-0.0498	0.0380	0.9514	-1.3091	0.1905	6.1
Phosphorous	-0.0574	0.0450	0.9442	-1.2768	0.2017	8.0
Magnesium	-0.0084	0.0283	0.9916	-0.2982	0.7655	2.1
Potassium	0.0047	0.0267	1.0047	0.1764	0.8600	2.6
Riboflavin	0.0432	0.0328	1.0441	1.3155	0.1884	3.1
Beta-carotene	-0.0328	0.0170	0.9677	-1.9263	0.0541	1.2
Total proteins	-0.0013	0.0288	0.9987	-0.0463	0.9630	3.1

*Also adjusted for total energy intake (kcal, continuous); smoking status (never, former, current); BMI <20, 20-22.9, 23-24.9, 25-29.9, 30-34.9, ≥35kg/m2); physical activity (inactive, moderately inactive, moderately active, active); diabetes history (no, yes); education status (none/primary, technical/professional, secondary, longer [including university]) and stratified by age at recruitment (<40, 40-44.9, 45-49.9, 50-54.9, 55-59.9, 60-64.9, 65-69.9, 70-74.9, ≥75), sex and recruitment centre.

VIF: Variance inflation factor. A value greater than 10 is indicative of multicollinearity.