

1 **Article type:** Original Research

2 **A prospective diet-wide association study for risk of colorectal cancer in EPIC**

3 Nikos Papadimitriou^{1,2*}, Emmanouil Bouras^{1*}, Piet A van den Brandt³, David C. Muller⁴, Areti
4 Papadopoulou¹, Alicia K Heath⁴, Elena Critselis⁵, Marc J Gunter², Paolo Vineis⁴, Pietro
5 Ferrari², Elisabete Weiderpass², Heiner Boeing⁶, Nadia Bastide⁷, Melissa A Merritt⁸, David S
6 Lopez^{9,10}, Manuela M Bergmann¹¹, Aurora Perez-Cornago¹², Matthias Schulze^{13,14}, Guri
7 Skeie¹⁵, Bernard Srour¹⁶, Anne Kirstine Eriksen¹⁷, Stina Boden¹⁸, Ingegerd Johansson¹⁹,
8 Therese Haugdahl Nøst¹⁵, Marco Lukic¹⁵, Fulvio Ricceri^{20,21}, Ulrika Ericson²², José María
9 Huerta^{23,24}, Christina C Dahm²⁵, Claudia Agnoli²⁶, Pilar Exezarreta Amiano^{27,24}, Anne
10 Tjønneland^{28,29}, Aurelio Barricarte Gurrea³⁰, Bas Bueno-de-Mesquita³¹, Eva Ardanaz^{24,32,33},
11 Jonna Berntsson³⁴, Maria-Jose Sánchez^{24,35,36,37}, Rosario Tumino³⁸, Salvatore Panico³⁹, Verena
12 Katzke¹⁶, Paula Jakszyn^{40,41}, Giovanna Masala⁴², Jeroen W.G. Derksen⁴³, J. Ramón Quirós⁴⁴,
13 Gianluca Severi^{45,46}, Amanda J Cross⁴, Ellio Riboli⁴, Ioanna Tzoulaki^{1,4}, Konstantinos K
14 Tsilidis^{1,4}

15 *These authors had equal contribution and are joint first

16 ¹Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina,
17 Greece

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

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

21 ⁴Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London,
22 London, UK

23 ⁵Biomedical Research Foundation of the Academy of Athens, Athens, Greece

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

25 ⁷Inserm, Centre for research in Epidemiology and Population Health (CESP), U1018, Nutrition,

26 Hormones and Women's Health team, Villejuif, F-94805, France.

27 ⁸University of Hawaii Cancer Center, Honolulu, Hawaii.

28 ⁹Department of Preventive Medicine and Population Health – UTMB School of Medicine, Galveston,

29 TX, USA.

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

31 ¹¹German Institute of Human Nutrition Potsdam-Rehbrücke, Germany

32 ¹²Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford,

33 United Kingdom

34 ¹³Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke,

35 Nuthetal, Germany

36 ¹⁴Institute of Nutrition Science, University of Potsdam, Nuthetal, Germany

37 ¹⁵Department of Community Medicine, Faculty of Health Sciences, University of Tromsø-The Arctic

38 University of Norway, Tromsø, Norway

39 ¹⁶Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

40 ¹⁷Danish Cancer Society Research Center, Diet, Genes and Environment, Copenhagen, Denmark

41 ¹⁸Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden

42 ¹⁹Department of Odontology, Umeå University, Sweden

43 ²⁰Department of Clinical and Biological Sciences, University of Turin, Italy

44 ²¹Unit of Epidemiology, Regional Health Service ASL TO3, Grugliasco (TO), Italy

45 ²²Department of Clinical Sciences in Malmö, Lund University, Malmö, Sweden

46 ²³Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain

47 ²⁴CIBER Epidemiology and Public Health CIBERESP, Madrid, Spain

48 ²⁵Department of Public Health, Aarhus University, Aarhus, Denmark

49 ²⁶Epidemiology and Prevention Unit Fondazione IRCCS Istituto Nazionale dei Tumori di Milano Via

50 Venezian, Milano, Italy

51 ²⁷Public Health Division of Gipuzkoa, BioDonostia Research Institute, Donostia-San Sebastian, Spain

- 52 ²⁸Danish Cancer Society Research Center, Diet, Genes and Environment, Copenhagen, Denmark
- 53 ²⁹Department of Public Health, University of Copenhagen, Copenhagen, Denmark
- 54 ³⁰Instituto de Salud Pública de Navarra, Navarra, Spain
- 55 ³¹Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the
56 Environment (RIVM), Bilthoven, the Netherlands
- 57 ³²Navarra Public Health Institute, Pamplona, Spain
- 58 ³³IdiSNA, Navarra Institute for Health Research, Pamplona, Spain
- 59 ³⁴Department of Clinical Sciences Lund, Oncology and Pathology, Lund University, Lund, Sweden
- 60 ³⁵Escuela Andaluza de Salud Pública (EASP), Granada, Spain
- 61 ³⁶Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain
- 62 ³⁷Department of Preventive Medicine and Public Health, University of Granada, Granada, Spain.
- 63 ³⁸Cancer Registry and Histopathology Department, Provincial Health Authority (ASP), Ragusa, Italy
- 64 ³⁹DIPARTIMENTO DI MEDICINA CLINICA E CHIRURGIA FEDERICO II UNIVERSITY,
65 NAPLES, ITALY
- 66 ⁴⁰Unit of Nutrition and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of
67 Oncology (ICO-IDIBELL), Barcelona, Spain
- 68 ⁴¹Blanquerna School of Health Sciences, Ramon Llull University, Barcelona, Spain
- 69 ⁴²Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and
70 Clinical Network - ISPRO, Florence, ITALY
- 71 ⁴³Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht
72 University, Utrecht, The Netherlands
- 73 ⁴⁴Public Health Directorate, Asturias, Spain
- 74 ⁴⁵Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, Équipe "Exposome et Hérité", CESP
75 UMR1018, 94805, Villejuif, France
- 76 ⁴⁶Department of Statistics, Computer Science and Applications (DISIA), University of Florence, Italy
- 77 **Corresponding author:**

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

81 **Word count: 4,240** (Introduction through discussion: 3,020 words; Figure/Table legends: 173
82 words; References: 1,047 words); **figures: 3; references: 35; supplementary tables: 11;**
83 **supplementary figures: 2; supplementary methods.**

84 **Disclosure:** The authors have declared no conflicts of interest.

Abstract (max 260 words, structured)

BACKGROUND & AIMS: Evidence regarding the association of dietary exposures with colorectal cancer (CRC) risk is not consistent with a few exceptions. Therefore, we conducted a diet-wide association study (DWAS) in the European Prospective Investigation into Cancer and Nutrition (EPIC) to evaluate the associations between several dietary exposures with CRC risk.

METHODS: The association of 92 food and nutrient intakes with CRC risk was assessed in 386,792 participants, 5,069 of whom developed incident CRC. 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.

RESULTS: 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.

CONCLUSIONS: 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

109 higher dietary intakes of phosphorus, magnesium, potassium, riboflavin, beta-carotene and
110 total protein.

111 **KEYWORDS:** nutrition; cohort study; colorectal cancer; epidemiology

112 **ABBREVIATIONS:** BMI: Body mass index; CI: Confidence interval; CRC: Colorectal
113 cancer; EPIC: European Prospective Investigation into Cancer and Nutrition; FDR: False
114 discovery rate; GWAS: Genome-wide association study; HR: Hazard ratio; MR: Mendelian
115 Randomization; NLCS: the Netherlands Cohort Study; DWAS: Diet-wide association study;
116 SNP: Single nucleotide polymorphism; WCRF: World Cancer Research Fund; WGS: Whole-
117 genome sequencing.

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 2018¹. 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 societies¹. 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 exceptions⁴. 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 risk⁴. Associations for other foods and nutrients and CRC risk exist, but are inconsistent and currently provide limited evidence according to WCRF⁴.

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 diet-wide association study (DWAS)⁵. The DWAS 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.

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 countries⁶. A total of 386,792 participants (71% women) were included in the present analysis after pertinent exclusions (supplementary methods).

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 that uses a case-cohort approach⁷. Of the 5,000 subcohort participants, 3,893 were included in the current analysis after pertinent exclusions (supplementary methods).

Assessment of dietary factors

In EPIC, consumption of foods over the last 12 months was assessed at baseline using validated country-specific food questionnaires⁶. In total, 92 dietary factors (63 foods and 29 nutrients) 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 reproducibility⁸ (supplementary methods).

Identification of colorectal cancer cases

In EPIC and NLCS, incident CRC cases were identified by record linkage with population-based cancer registries or a combination of registries, insurance records and active follow up of the study participants or their relatives. More details are provided in the supplementary methods.

Statistical analyses

In EPIC, separate Cox proportional hazards regression models were used to investigate the associations between each of the dietary factors with CRC risk. In NLCS, given the case-cohort design, Prentice weighted Cox proportional hazards regression models with robust standard error estimation were implemented⁹. All of the models were adjusted for: total energy intake; smoking; body mass index (BMI); physical activity; diabetes history; level of education

and family history of CRC (in NLCS only) and further stratified by sex, age and in EPIC also by centre.

To account for multiple comparisons, the false discovery rate (FDR) adjusted *P* values (or *q* values) were estimated for each association analysed¹⁰. 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. Presence of non-linear associations was investigated using restricted cubic spline models. More details on the statistical analyses methods are provided in the supplementary methods.

Results

Study characteristics

After a mean follow up of 14.1 (SD:3.9) 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) 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 **Supplementary 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%), 47% were overweight or obese, one third (33%) were never smokers, and 48% spent more than 60 minutes/day on non-occupational physical activities.

DWAS 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 2**). Higher intakes of alcohol, liquor/spirits, wine, beer/cider, soft drinks, and pork were positively associated with CRC, whereas higher milk, cheese, calcium, phosphorus, magnesium, potassium, riboflavin, vitamin B6, beta-carotene, fruit, fibre, non-white bread, banana, and total protein 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, with associations being inverse for colon cancer and null for rectal cancer (**Supplementary Table 3**). Regarding proximal versus distal colon subsites, only total alcohol and wine had heterogeneous results, whereby the associations were positive only for distal colon cancer (**Supplementary Table 4**). Additionally, heterogeneous associations were observed by sex, for total alcohol and spirits, magnesium, fibre, and non-white bread, where the associations were only observed in men (**Supplementary Table 5**). 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**). There was some evidence for non-linearity (P value = 0.028) in the association of alcohol intake and CRC risk (**Supplementary Figure 2**).

Replication analysis in NLCS

Of the 20 associations with an FDR<0.05 in EPIC, four associations reached nominal statistical significance (P value < 0.05) in the NLCS cohort in the analysis for CRC (**Figure 2; Supplementary Table 6**), namely alcohol and liquor/spirits (positively), milk and calcium intake (inversely). An additional four associations, namely phosphorus, magnesium, riboflavin

and total protein, had a borderline inverse association in NLCS, 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, potassium, vitamin B6, beta-carotene and total protein in the analysis for colon versus rectal cancer (Supplementary Tables 7-8). Little heterogeneity was observed by sex for CRC risk (Supplementary Table 9).

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. The remaining 13 associations yielded a nominally significant summary finding: alcohol (HR per 1 SD increment in intake per day: 1.07; 95% CIs: 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) (*Figure 2*; Supplementary Table 6).

Pairwise correlations and Mutual-adjustment analysis

Most of the pair-wise correlation coefficients for the 20 FDR-significant foods/nutrients in EPIC were weak and ranged from -0.25 to 0.79 (*Figure 3*).

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 (1.05; 1.03-1.11) **Supplementary Table 10).**

Gene-Nutrient interaction analysis

Of the 73×20 gene-nutrient multiplicative interactions that were tested, using the Bonferroni adjusted *P* value threshold of 3.4×10^{-5} , no interaction remained significant (**Supplementary Table 11).**

Discussion

We used the DWAS 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 strong¹¹. Persons with higher total alcohol consumption had a higher risk of CRC (summary HR per SD increment in intake/day: 1.07; 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^{12, 13}. Acetaldehyde, as a metabolite of ethanol oxidation, can be carcinogenic in colonocytes¹⁴. Mendelian randomization (MR)

studies have failed to demonstrate an association between genetically proxied alcohol consumption and CRC risk, but this analysis was underpowered to detect relatively small effects¹⁵.

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 colorectum¹⁶. Heterogeneity by anatomical subsite or sex was not observed, in agreement with the WCRF meta-analysis and a more recent publication in the Nurses' Health Study^{11, 12}. 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 risk¹⁷. 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 effects¹⁸. Evidence from MR studies suggests that genetically proxied milk consumption is associated with a reduced CRC risk, but failed to demonstrate an association for genetically proxied calcium or phosphorus concentrations^{19, 20}. Additionally, although previous RCTs have showed null associations for calcium supplementation in relation to CRC risk, a 13% decreased risk of colorectal adenoma recurrence has been reported in a meta-analysis of four RCTs, with daily doses of calcium ranging from 1200 to 2000 mg²¹.

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 gram/day increment for red meat and 16% per 50 gram/day for processed meat⁴. 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 studies²². Red meat also contains haem iron at high levels that may stimulate the endogenous formation of carcinogenic N-nitroso compounds, which promote colorectal tumourigenesis²³. 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; 0.98-1.05; HR per 31.5 grams of processed meat intake daily: 1.04; 1.00-1.08). 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 risk²⁴. 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, 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 mortality²⁵.

The current DWAS 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 studies²⁶. 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 cells²⁷. Furthermore, magnesium has been shown to improve insulin sensitivity and lower plasma insulin concentrations, which may have an impact on CRC development²⁸. No association of genetically proxied circulating concentrations of **magnesium** was found in a recent MR study²⁰

We also observed an inverse association between intake of **beta-carotene** and risk of CRC, but few other studies have investigated this association²⁹. Our findings agree with a previous report from EPIC in 2014²⁹. 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 CRC³⁰. No evidence of association between genetically proxied circulating concentrations of **beta-carotene** and CRC have been reported in MR studies²⁰.

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 methylation³¹. 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 scarce³². 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 considered³². 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 risk³³.

Little is known on the role that **potassium** may play in relation to CRC risk, and epidemiological evidence thus far is limited³⁴. 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.

Strengths of this study include its large size and long follow-up duration and the DWAS 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. Intercorrelations between dietary exposures may have led to low precision of the estimates, even though variance inflation factors were relatively small, which might explain that none of the dietary factors remained in the multivariable adjusted model. Furthermore, it is possible that there might be an association for foods or nutrients that were not included in this analysis. Additionally, the data derived from the Dutch food composition table were not checked against the use of ENDB for nutrient calculation, so discrepancies may have occurred, hence it is possible that some of the discrepancies observed between the two cohorts for some dietary exposures, are in part due to poor reproducibility in measurements. Among the exposures for which heterogeneity was observed between EPIC and NLCS, correlation

between the baseline FFQs and 24-hour diet recalls has been reported to be good for fruit, fibre, vitamin B6 and beverage consumption in NLCS and fairly good for fibre and fruit across most EPIC centres, but for exposures like non-white bread or vitamin B6 no information was available^{8, 35}. 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.

369 **Acknowledgement**

370 We are grateful to all the participants who have been part of the project and to the EPIC study
371 team at the University of Cambridge who has enabled this research.

372 **Funding sources**

373 This work was supported by the World Cancer Research Fund International Regular Grant
374 Programme (WCRF 2014/1180 to Konstantinos K. Tsilidis).

375 The coordination of EPIC is financially supported by International Agency for Research on
376 Cancer (IARC) and also by the Department of Epidemiology and Biostatistics, School of Public
377 Health, Imperial College London which has additional infrastructure support provided by the
378 NIHR Imperial Biomedical Research Centre (BRC).

379 The national cohorts are supported by: Danish Cancer Society (Denmark); Ligue Contre le
380 Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Éducation Nationale, Institut National
381 de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German
382 Cancer Research Center (DKFZ), German Institute of Human Nutrition Potsdam-Rehbruecke
383 (DIfE), Federal Ministry of Education and Research (BMBF) (Germany); Associazione
384 Italiana per la Ricerca sul Cancro-AIRC-Italy, Compagnia di SanPaolo and National Research
385 Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands
386 Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg
387 Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The
388 Netherlands); Health Research Fund (FIS) - Instituto de Salud Carlos III (ISCIII), Regional
389 Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan
390 Institute of Oncology - ICO (Spain); Swedish Cancer Society, Swedish Research Council and
391 County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-
392 Norfolk; C8221/A29017 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-
393 Norfolk; MR/M012190/1 to EPIC-Oxford) (United Kingdom).

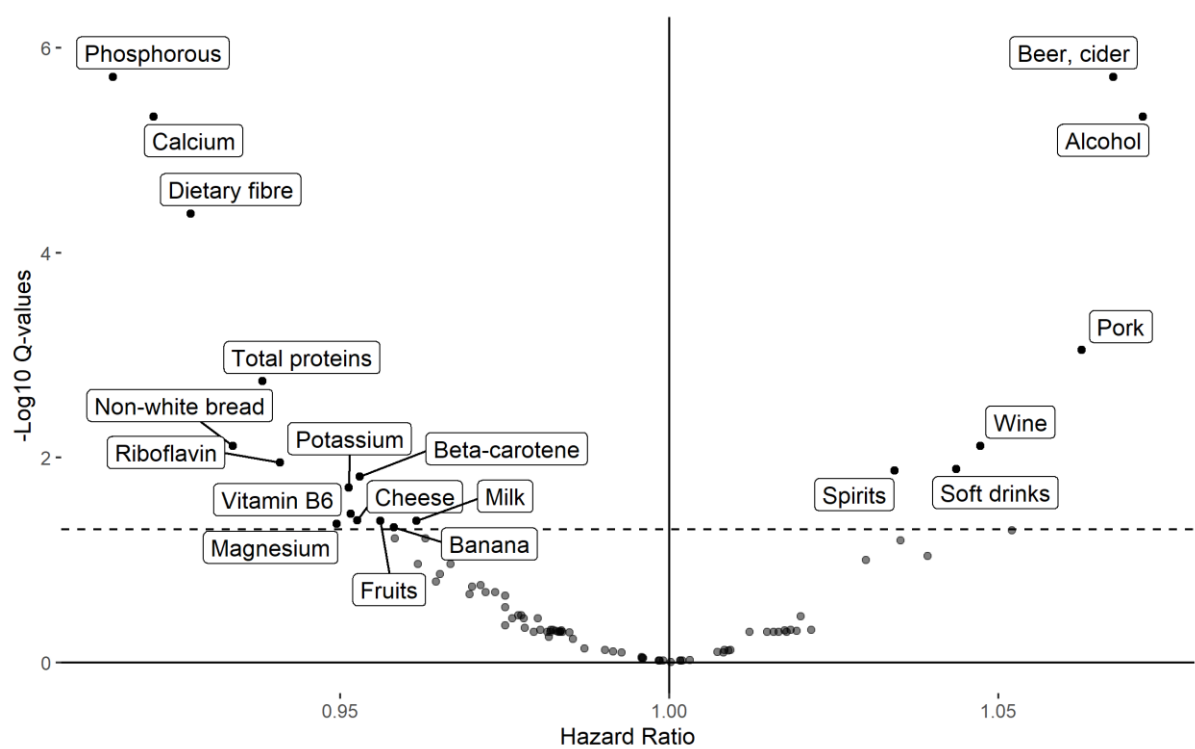
394 **Disclaimer:** Where authors are identified as personnel of the International Agency for
395 Research on Cancer / World Health Organization, the authors alone are responsible for the
396 views expressed in this article and they do not necessarily represent the decisions, policy or
397 views of the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

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

401 **References**

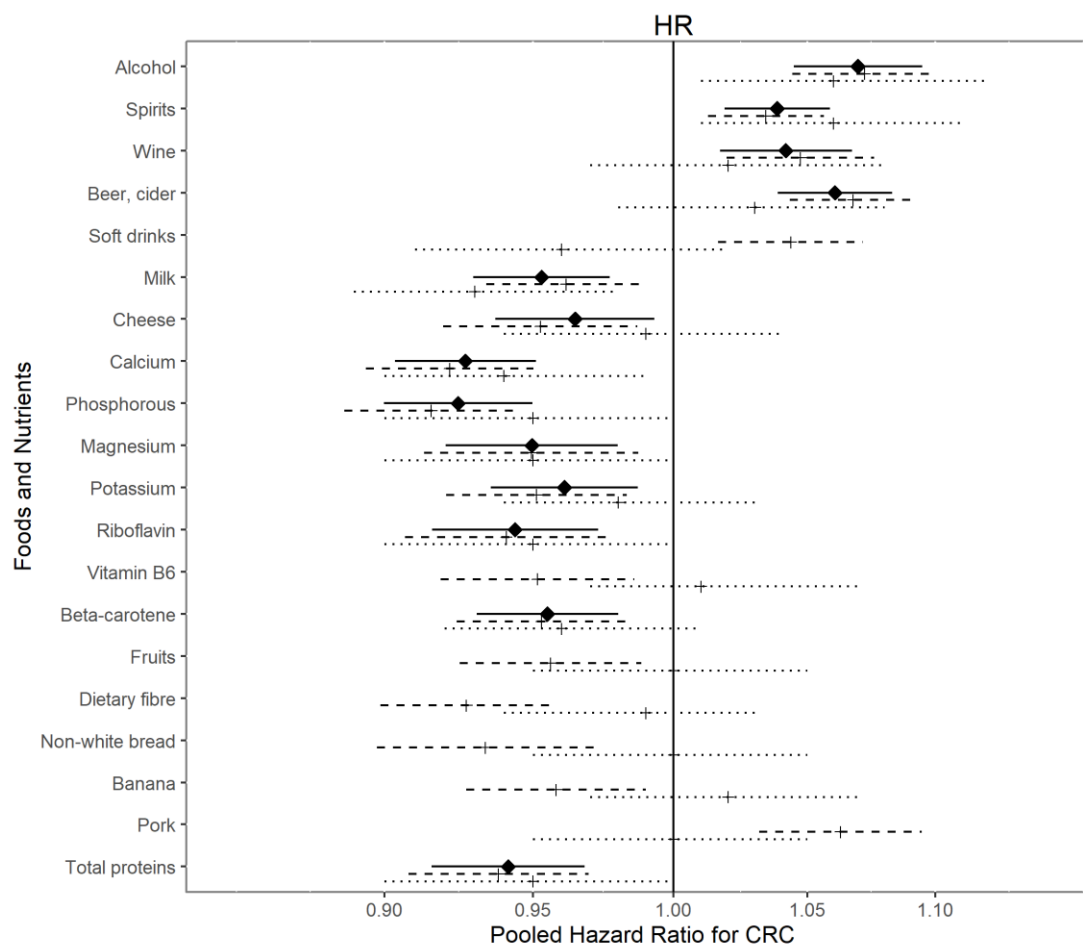
- 402 1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN
403 estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer
404 journal for clinicians. 2018;68(6):394-424. Epub 2018/09/13.
- 405 2. Kyrgiou M, Kalliala I, Markozannes G, et al. Adiposity and cancer at major anatomical
406 sites: umbrella review of the literature. BMJ (Clinical research ed). 2017;356:j477. Epub
407 2017/03/02.
- 408 3. Rezende LFM, Sa TH, Markozannes G, et al. Physical activity and cancer: an umbrella
409 review of the literature including 22 major anatomical sites and 770 000 cancer cases. British
410 journal of sports medicine. 2018;52(13):826-33. Epub 2017/11/18.
- 411 4. World Cancer Research Fund, American Institute for Cancer research. Food,
412 Nutrition,Physical Activity, and the Prevention of Colorectal Cancer. Continuous Update
413 Project Report. 2018.
- 414 5. Papadimitriou N, Muller D, van den Brandt PA, et al. A nutrient-wide association study
415 for risk of prostate cancer in the European Prospective Investigation into Cancer and Nutrition
416 and the Netherlands Cohort Study. European journal of nutrition. 2019. Epub 2019/11/11.
- 417 6. Riboli E, Hunt KJ, Slimani N, et al. European Prospective Investigation into Cancer
418 and Nutrition (EPIC): study populations and data collection. Public Health Nutr.
419 2002;5(6b):1113-24. Epub 2003/03/18.
- 420 7. van den Brandt PA, Goldbohm RA, van 't Veer P, et al. A large-scale prospective cohort
421 study on diet and cancer in The Netherlands. J Clin Epidemiol. 1990;43(3):285-95.
- 422 8. Goldbohm RA, van den Brandt PA, Brants HA, et al. Validation of a dietary
423 questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin
424 Nutr. 1994;48(4):253-65.
- 425 9. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease
426 prevention trials. Biometrika. 1986;73(1):1-11.
- 427 10. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful
428 approach to multiple testing. Journal of the Royal Statistical Society. 1995;Ser B(57):289-300.
- 429 11. World Cancer Research Fund/American Institute for Cancer Research. Continuous
430 Update Project Expert Report 2018. Diet, nutrition, physical activity and colorectal cancer.
431 Available at dietandcancerreport.org.
- 432 12. Wei EK, Colditz GA, Giovannucci EL, et al. A Comprehensive Model of Colorectal
433 Cancer by Risk Factor Status and Subsite Using Data From the Nurses' Health Study.
434 American journal of epidemiology. 2017;185(3):224-37. Epub 2017/01/12.
- 435 13. Murphy N, Ward HA, Jenab M, et al. Heterogeneity of Colorectal Cancer Risk Factors
436 by Anatomical Subsite in 10 European Countries: A Multinational Cohort Study. Clinical
437 gastroenterology and hepatology : the official clinical practice journal of the American
438 Gastroenterological Association. 2019;17(7):1323-31.e6. Epub 2018/07/30.
- 439 14. Seitz HK, Stickel F. Molecular mechanisms of alcohol-mediated carcinogenesis.
440 Nature reviews Cancer. 2007;7(8):599-612. Epub 2007/07/25.
- 441 15. Cornish AJ, Law PJ, Timofeeva M, et al. Modifiable pathways for colorectal cancer: a
442 mendelian randomisation analysis. Lancet Gastroenterol Hepatol. 2020;5(1):55-62. Epub
443 10/24.
- 444 16. Newmark HL, Wargovich MJ, Bruce WR. Colon cancer and dietary fat, phosphate, and
445 calcium: a hypothesis. J Natl Cancer Inst. 1984;72(6):1323-5. Epub 1984/06/01.
- 446 17. Moskal A, Freisling H, Byrnes G, et al. Main nutrient patterns and colorectal cancer
447 risk in the European Prospective Investigation into Cancer and Nutrition study. British journal
448 of cancer. 2016;115(11):1430-40. Epub 2016/10/21.
- 449 18. Welch AA, Fransen H, Jenab M, et al. Variation in intakes of calcium, phosphorus,
450 magnesium, iron and potassium in 10 countries in the European Prospective Investigation into
451 Cancer and Nutrition study. Eur J Clin Nutr. 2009;63 Suppl 4:S101-21. Epub 2009/11/06.

19. Larsson SC, Mason AM, Kar S, et al. Genetically proxied milk consumption and risk of colorectal, bladder, breast, and prostate cancer: a two-sample Mendelian randomization study. *BMC Med.* 2020;18(1):370-.
20. Tsilidis KK, Papadimitriou N, Dimou N, et al. Genetically predicted circulating concentrations of micronutrients and risk of colorectal cancer among individuals of European descent: a Mendelian randomization study. *Am J Clin Nutr.* 2021. Epub 2021/03/20.
21. Bonovas S, Fiorino G, Lytras T, et al. Calcium supplementation for the prevention of colorectal adenomas: A systematic review and meta-analysis of randomized controlled trials. *World J Gastroenterol.* 2016;22(18):4594-603.
22. Cross AJ, Sinha R. Meat-related mutagens/carcinogens in the etiology of colorectal cancer. *Environmental and molecular mutagenesis.* 2004;44(1):44-55. Epub 2004/06/17.
23. Santarelli RL, Pierre F, Corpet DE. Processed meat and colorectal cancer: a review of epidemiologic and experimental evidence. *Nutrition and cancer.* 2008;60(2):131-44. Epub 2008/04/30.
24. Norat T, Bingham S, Ferrari P, et al. Meat, fish, and colorectal cancer risk: the European Prospective Investigation into cancer and nutrition. *J Natl Cancer Inst.* 2005;97(12):906-16. Epub 2005/06/16.
25. Zheng Y, Li Y, Satija A, et al. 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. Epub 2019/06/14.
26. Meng Y, Sun J, Yu J, et al. Dietary Intakes of Calcium, Iron, Magnesium, and Potassium Elements and the Risk of Colorectal Cancer: a Meta-Analysis. 2019;189(2):325-35.
27. Mori H, Morishita Y, Shinoda T, et al. Preventive effect of magnesium hydroxide on carcinogen-induced large bowel carcinogenesis in rats. *Basic life sciences.* 1993;61:111-8. Epub 1993/01/01.
28. van den Brandt PA, Smits KM, Goldbohm RA, et al. Magnesium intake and colorectal cancer risk in the Netherlands Cohort Study. *British journal of cancer.* 2007;96(3):510-3. Epub 2007/02/08.
29. Leenders M, Leufkens AM, Siersema PD, 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(12):2930-9. Epub 2014/04/29.
30. Malila N, Virtamo J, Virtanen M, et al. Dietary and serum alpha-tocopherol, beta-carotene and retinol, and risk for colorectal cancer in male smokers. *Eur J Clin Nutr.* 2002;56(7):615-21. Epub 2002/06/25.
31. Newman AC, Maddocks ODK. One-carbon metabolism in cancer. *British journal of cancer.* 2017;116(12):1499-504. Epub 05/04.
32. Zschäbitz S, Cheng T-YD, Neuhaus ML, et al. 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(2):332-43. Epub 12/19.
33. Larsson SC, Orsini N, Wolk A. Vitamin B6 and risk of colorectal cancer: a meta-analysis of prospective studies. *Jama.* 2010;303(11):1077-83. Epub 2010/03/18.
34. Key TJ, Appleby PN, Masset G, et al. Vitamins, minerals, essential fatty acids and colorectal cancer risk in the United Kingdom Dietary Cohort Consortium. *International journal of cancer.* 2012;131(3):E320-E5.
35. 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. Epub 1997/01/01.



504

505 **Figure 1. Volcano plot showing results from the DWAS regarding the association between**
506 **92 dietary factors and CRC risk in EPIC.** The Y-axis shows the false discovery rate (FDR)
507 adjusted P values in $-\log_{10}$ scale from the Cox proportional hazards models for each dietary
508 factor. The X-axis shows the estimated HR for each dietary factor per 1 SD increase in daily
509 consumption. The dashed horizontal line represents the level of significance corresponding to
510 FDR of 5%.



511

512 **Figure 2. Forest plot showing the hazard ratios and 95% confidence intervals for the 20**

513 **FDR significant associations (FDR<5%), in EPIC (---) and NLCS (···), as well as the**

514 **results from a meta-analysis (MA) (—). The X-axis shows the estimated HR for each dietary**

515 **factor for 1 SD increase in daily consumption. The diamond and the solid line represent the**

516 **pooled HR and 95%CI of the MA. MA was not performed when heterogeneity was high.**

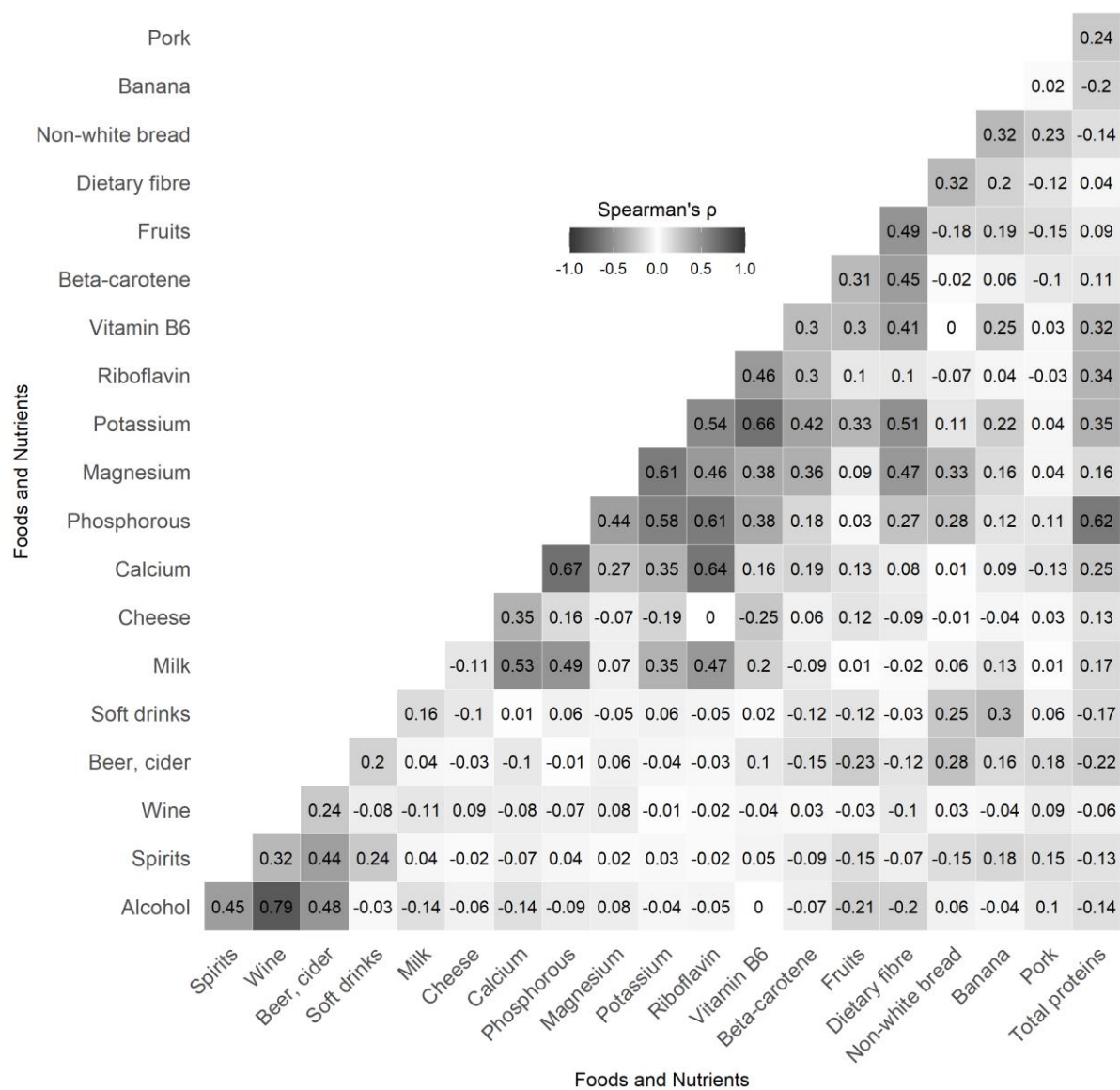


Figure 3. Pairwise partial correlation coefficients (Spearman's ρ) of the 20 FDR-significant foods/nutrients in EPIC, adjusting for age, sex and centre.