The 3V score and joint associations of low ultra-processed food, biodiverse and plant-based diets on colorectal cancer risk: results from the European Prospective Investigation into Cancer and Nutrition (EPIC) study

Emine Koc Cakmak,[a,b,*] Aline Al Nahas,[c] Bernadette Chimera,[c] Giles Hanley-Cook,[d] Jeroen Berden,[c,d] Anthony Fardet,[e] Edmond Rock,[e] Carine Biessy,[c] Geneviève Nicolas,[c] Nathalie Kliemann,[f] Fernanda Rauber,[g,h] Renata Bertazzi Levy,[g,h,i] Lorenzo Mangone,[a] Mathilde Touvier,[j] Bernard Srour,[j] Emmanuelle Kesse-Guyot,[j] Carl Lachat,[d] Guri Skeie,[k] Elisabete Weiderpass,[l] Franziska Jannasch,[m] Christina C. Dahm,[n] Daniel Borch Ibsen,[n,o] Christina Dahl,[p] Cecilie Kyrø,[p] Mariem Hajji-Louati,[q] Chloé Marques,[q] Gianluca Severi,[q,r] Verena Katzke,[s] Rudolf Kaaks,[s] Matthias B. Schulze,[m,n,o,p,q,r,s,t] Saverio Caini,[u] Sabina Sieri,[v] Maria Santucci De Magistris,[w] Rosario Tumino,[x] Carlotta Sacerdote,[y,z] Raúl Zamora-Ros,[aa] Maria-José Sánchez,[ab,ac,ad] Ana Jimenez-Zabala,[ae,af] Jesús-Humberto Gómez,[ag,ah] Marcela Guevara,[ai,aj,ak] Elio Riboli,[a] Marc J. Gunter,[a] Inge Huybrechts,[c] Paolo Vineis,[a,b] and Oliver J. K. Robinson[a,b]

[a]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom
[b]Medical Research Council Centre for Environment and Health, Imperial College London, London, United Kingdom
[c]Nutrition and Metabolism Branch, International Agency for Research on Cancer, Lyon, France
[d]Department of Food Technology, Safety and Health, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium
[e]INRAE, Université Clermont Auvergne, UNH, Human Nutrition Unit, CRNH Auvergne, F-63000, Clermont-Ferrand, France
[f]Centro de Pesquisas Oncológicas (CEPON), Santa Catarina, Brazil
[g]Center for Epidemiological Research in Nutrition and Health, University of São Paulo, Av. Dr Arnaldo 715, São Paulo, SP, 01246-904, Brazil
[h]Department of Preventive Medicine, School of Medicine, University of Sao Paulo, Av. Dr. Arnaldo, 455, Sao Paulo, SP, 01246-903, Brazil
[i]University of Salamanca, Institute of Biomedical Research of Salamanca (IBSAL), Paseo de San Vicente, 58-182, 37007, Salamanca, Spain
[j]Université Sorbonne Paris Nord and Université Paris Cité, INSERM, INRAE, CNAM, Centre of Research in Epidemiology and StatisticS (CRESS), Nutritional Epidemiology Research Team (EREN), F-93017, Bobigny, France
[k]Department of Community Medicine, UIT The Arctic University of Norway, Tromsø, Norway
[l]International Agency for Research on Cancer, Lyon, France
[m]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur Scheunert-Allee 114 –116, 14558, Nuthetal, Germany
[n]Department of Public Health, Aarhus University, Aarhus, Denmark
[o]Steno Diabetes Center Aarhus, Aarhus, Denmark
[p]Diet, Cancer and Health, Danish Cancer Institute, Copenhagen, Denmark
[q]Paris-Saclay University, UVSQ, Inserm, Gustave Roussy, CESP, Villejuif, France
[r]Department of Statistics, Computer Science and Applications "G. Parenti", University of Florence, Italy
[s]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[t]Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany
[u]Cancer Risk Factors and Lifestyle Epidemiology Unit, Institute for Cancer Research, Prevention, and Clinical Network (ISPRO), Florence, Italy
[v]Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale Dei Tumori Di Milano, Milan, Italy
[w]Department of Clinical and Experimental Medicine, Federico II University, Naples, Italy
[x]Hyblean Association for Epidemiology Research AIRE ONLUS Ragusa, Italy
[y]Department of Health Sciences, University of Eastern Piedmont, Via Solaroli 17, 28100, Novara, Italy
[z]Unit of Epidemiology, Local Health Unit of Novara, Viale Roma, 7, 128100, Novara, Italy
[aa]Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology, Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain
[ab]Escuela Andaluza de Salud Pública (EASP), 18011, Granada, Spain
[ac]Instituto de Investigación Biosanitaria ibs.GRANADA, 18012, Granada, Spain
[ad]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), 28029, Madrid, Spain
[ae]Ministry of Health of the Basque Government, Sub-Directorate for Public Health and Addictions of Gipuzkoa, San Sebastián, Gipuzkoa, Spain

*Corresponding author. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom.
E-mail address: e.koc21@imperial.ac.uk (E. Koc Cakmak).

[af]Biogipuzkoa Health Research Institute, Group of Epidemiology of Chronic and Communicable Diseases, San Sebastián, Gipuzkoa, Spain
[ag]Department of Epidemiology, Regional Health Council, IMIB-Arrixaca, Murcia, Spain
[ah]CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[ai]Instituto de Salud Pública y Laboral de Navarra, 31003, Pamplona, Spain
[aj]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), 28029, Madrid, Spain
[ak]Navarra Institute for Health Research (IdiSNA), 31008, Pamplona, Spain

eClinicalMedicine
2025;90: 103662
Published Online 1 December 2025
https://doi.org/10.1016/j.eclinm.2025.103662

Summary

Background Diet may modify colorectal cancer risk. We investigated the associations of three dietary patterns, ultra-processed food (UPF) consumption, healthy plant-based food consumption, and food biodiversity, separately and combined into a "3V" score with risk of colorectal cancer.

Methods This study used data from the prospective European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, which recruited participants between 1992, and 2000, from 23 centres in ten European countries. The 3V score was developed by standardising and summing the healthy plant diet index (hPDI) and dietary species richness per year (DSR) and subtracting UPF (Nova category 4) intake in % g/day. Associations with colorectal cancer risk were assessed among 450,111 middle-aged participants of the EPIC cohort using multivariable-adjusted Cox regression models. Independent associations of each 3V component were assessed using mutually adjusted models. Data-driven thresholds were applied to assess adherence to the 3V components, set at the minimum value of the fourth quintile for hPDI, DSR and low UPF.

Findings During mean (standard deviation (SD)) follow-up of 14.9 (4) years, absolute colorectal cancer rates were 8.59 and 10.37 cases/10,000 person-years for the highest and lowest quintiles of the 3V score, respectively. Inverse associations were found for colorectal (hazard ratio (HR) comparing highest vs lowest quintile: 0.84; 95% confidence interval (CI): 0.76–0.94), colon (HR: 0.82; 95% CI: 0.72–0.93), and distal colon cancer (HR: 0.81; 95% CI: 0.67–0.99), with significant linear trends observed across quintiles. UPF intake was positively associated with colon cancer risk (HR per 1 SD increment: 1.06; 95% CI: 1.02–1.11) when mutually adjusted for the other 3V components. Adherence to low UPF, high hPDI, and high DSR was inversely associated with colorectal (HR: 0.73; 95% CI: 0.61–0.88), colon (HR: 0.72; 95% CI: 0.57–0.91), and rectal cancer (HR: 0.65; 95% CI: 0.46–0.91) compared to adhering to none.

Interpretation Adherence to the 3V diet is associated with lower risk of colorectal cancers.

Funding Cancer Research UK, World Cancer Research Fund.

Keywords: Food processing; Food biodiversity; Prospective cohort; Dietary risk factors; Colorectal cancer

Introduction

Colorectal cancer is one of the most frequently occurring malignancies worldwide, with over 1.9 million new cases and 935,000 deaths reported in 2020.[1] Globally, it ranks third in incidence among cancers affecting both sexes, and it is the second leading cause of cancer-related mortality in females and third in males.[1]

Colorectal cancer is a heterogeneous disease; indeed, colorectal tumours that have developed at distinct anatomical sites have dissimilar molecular[2] and clinical characteristics.[3] Research suggests that tumours located at different anatomical sites potentially have diverse aetiologies and risk factors.[4] Although colorectal cancer is a complex multifactorial disease, the majority of colorectal cancer cases are likely preventable[5] and amenable to public health strategies supporting healthy lifestyle habits.[3] Diet is considered a major contributor to colorectal carcinogenesis, involving both protective and risk factors.[6]

Recent research has indicated that dietary factors such as ultra-processed food (UPF) consumption, plant-based diets, and food biodiversity may influence cancer risk. UPFs are made with industrial formulations of ingredients that include protein isolates, modified oils, and other food substances of no culinary use, along with flavours, colours, and other additives designed to make the product intensely attractive to consumers. Plant-based diet indices (PDIs) quantify dietary patterns based on consumption of healthy and less healthy plant foods and animal foods. Specifically, PDI, healthful PDI (hPDI), and unhealthful PDI (uPDI) assign negative scores to animal foods and vary scores for plant foods

Research in context

Evidence before this study

Some studies suggested associations between higher intakes of UPFs and greater risk of colorectal cancer, while more biodiverse and plant-based diets have been associated with lower gastrointestinal cancer risks. Though these dietary risk factors have been explored in isolation, their potential interactions or joint relationships with the risk of colorectal cancer have not yet been examined.

Added value of this study

Using the 3V framework (**V**rai/Veritable, **V**égétal/Vegetal-based, and **V**arié/Varied), this study is the first to underscore the potential combined benefits of adopting low ultra-processed, healthy plant-based and biodiverse dietary patterns. Higher 3V score was inversely associated with overall colorectal cancer risk and, in particular, colon and distal colon cancers within a pan-European cohort. When the independent association of each 3V component with cancer risk at anatomical subsites were assessed, UPF intake was identified as the primary driver of the positive association with colon cancer risk, while the other components of the 3V framework, including the plant-based diet and food biodiversity, showed weaker or non-significant associations. Adherence to all three pillars of the cumulative 3V score — low ultra-processed food, more nutritionally healthful plant-based food, and greater food biodiversity—was associated with significant reductions in cancer risk over ~15 years of follow-up: 27% for colorectal cancer (95% CI: 0.61–0.88), 28% for colon cancer (95% CI: 0.57–0.91), and 35% for rectal cancer (95% CI: 0.46–0.91), as compared to those participants who adhered to none.

Implications of all the available evidence

This study highlights the importance of integrating the three complementary dietary dimensions in colorectal cancer aetiology and preventive nutrition. These dietary dimensions are conceptually linked to broader sustainability issues, including biodiversity loss. As such, public health policies and programs should facilitate reductions in ultra-processed food consumption, promoting healthy plant-based diets and increasing food biodiversity to lower colorectal cancer risk.

based on their nutritional quality. Food biodiversity, captured through dietary species richness (DSR), reflects the number of distinct taxonomic species consumed in the diet.

Observational studies have reported associations between higher UPF consumption and higher risk of colorectal cancer.[7–10] In the EPIC cohort, Al Nahas et al. reported a positive association between higher UPF consumption and greater colorectal cancer risk.[11] Conversely, PDIs, in particular nutritionally hPDI, have been related to a lower risk of colorectal cancer, while uPDI has been associated with higher risk.[12,13] Furthermore, Kim et al. reported that greater adherence to various healthy plant-based diets was associated with lower colorectal cancer risk in men.[14] Moreover, prospective studies from the EPIC cohort have shown that greater food biodiversity is associated with lower overall mortality, cause-specific mortality due to cancer, and gastrointestinal cancer risk in Europe.[15,16]

The 3V diet—**V**rai (real, "véritable", low UPF), **V**égétal (vegetal-based), and **V**arié ("Varied", biodiverse, preferably consisting of "organic, local, seasonal" foods)—proposed by Fardet and Rock,[17] offers a practical, simple and comprehensive set of dietary recommendations that integrate three pillars of a healthy diet rooted in sustainable agrifood systems. The 3V score was developed to jointly assess plant-based diet quality, food processing, and food biodiversity measured through dietary species richness—three dimensions increasingly prioritised in diet–health–sustainability research. This approach emphasises first reducing UPF consumption, then increasing the intake of nutritionally healthful plant-based foods and finally promoting dietary biodiversity.

We investigated the relationship between the 3V diet and its components and colorectal cancer risk, including the various anatomical subsites, in a large cohort across nine European countries. Our approach aims to elucidate how adherence to the Vrai, Végétal, and Varié principles collectively and independently relate to colorectal cancer risk to enhance our understanding and propose effective diet-related cancer prevention strategies. These three dietary components were selected due to their ability to capture distinct but complementary aspects of diet related to food composition, food processing, and food biodiversity—factors that are especially relevant for colorectal cancer risk and sustainable dietary recommendations.

Methods

Ethics

This study adheres to the Declaration of Helsinki. The EPIC study was approved by the International Agency for Research on Cancer (IARC) Ethics Committee (IEC) (Ref# IEC 22-36) and local ethics committees of the study centres. All participants provided written informed consent for the collection, storage, and individual follow-up of their data.

Study participants

The European Prospective Investigation into Cancer and Nutrition (EPIC) is a cohort comprised of 521,324 participants (70.1% female) who were recruited

between 1992 and 2000 across 23 centres in ten European countries (Denmark, France, Germany, Italy, the Netherlands, Norway, Spain, Sweden, Greece, and the United Kingdom).[18] The EPIC cohort was designed to investigate the relationships between diet, nutrition, environmental factors, and cancer and other chronic disease incidences. The current study includes data from 450,111 middle-aged participants. Exclusion criteria are outlined in Supplementary Figure S1. EPIC-Greece data were not available for this analysis.

Baseline characteristics

In EPIC, centralised country-specific dietary food frequency questionnaires (FFQ) were administered at baseline, either as self-administered or interviewer-administered semi-quantitative FFQs, depending on the centre; Malmö (Sweden) used a non-quantitative FFQ with a 7-day diet record[19] (Supplementary Table S1). Additionally, data on physical activity, education, smoking, and alcohol intake was collected using questionnaires at the baseline, as described in detail elsewhere.[18] Body weight and height were measured at all centres, except for France, Norway, and Oxford, where participants provided self-reported measurements that were subsequently verified.[18]

Dietary exposures

Level of processing using the Nova classification

UPFs (% g/day) were defined according to the Nova food processing classification system.[20] UPF consumption was expressed as the proportion (%) of total weight (g/day) rather than energy, to account for non-caloric UPFs (e.g., artificially sweetened beverages) and food processing factors not captured by energy-based measures. This method provides a more comprehensive assessment of UPF consumption. The Nova classification includes four categories. Nova 1 includes unprocessed or minimally processed foods, which are natural foods or natural foods altered by methods such as freezing, pasteurisation, and other processes that do not add additional salt, sugar, oils, fats, or other food substances such as fruits and vegetables with no added sugars or substances. Nova 2 includes processed culinary ingredients derived from foods which are unprocessed/minimally processed (e.g. oil, butter) or natural (e.g. salt). Nova 3 includes processed foods, which are products (industrially produced or not) created by combining unprocessed/minimally processed foods (Nova 1) and processed culinary ingredients (Nova 2) through various preservation techniques, including canning and bottling. Examples of foods in Nova 3 are bread, cheese, beer, wine, and smoked fish. Lastly, Nova 4 refers to UPFs, which are made with industrial formulations of ingredients that, besides sugar, oil, or salt, include protein isolates, modified oils, and other food substances of no culinary use, along with flavours, colours, and other additives designed to make the product intensely attractive to consumers. Examples of foods in this group are processed meats (e.g. reconstituted meat products or sausage), (carbonated) soft drinks, packaged long shelf bread, as well as sweet or savoury packaged snacks, dehydrated soups, flavoured yoghurts, and ready-to-eat meals. Further details on the UPF subgroups are available in Supplementary Table S2. In EPIC, three scenarios (lower, middle, and upper bound) were applied to address the uncertainty in the classification of foods such as bread that may have been consumed as processed or UPF at baseline.[19] Using biomarkers of industrial processing (elaidic acid and 4-methylsyringol sulphate) for validation of UPF, the middle-bound scenario showed the best classification performance[19] and was therefore used for subsequent analyses.

Plant-based diet scores

Using data on plant-based food group intakes (g/day) from the EPIC FFQs, we calculated total PDI, hPDI, and uPDI, based on 18 food groups following a previously described methodology.[21] The food groups were categorised into healthy plant-based foods (e.g. whole grains, fruits, vegetables, nuts, legumes, vegetable oils, and tea/coffee), less healthy plant-based foods (including fruit juices, refined grains, potatoes, sugar-sweetened beverages, sweets, and desserts), and animal-source foods (e.g. animal fat, dairy, eggs, fish/seafood, meat, and other miscellaneous animal products). Further details on the 18 plant-based food groups are available in Supplementary Table S3.

For each food group across the three indices, daily consumption in grams was divided into fifths. In the PDI, all plant-based food groups were positively scored, with the lowest fifth of consumption receiving 1 point and the highest fifth receiving 5 points, while animal food groups were inversely scored, with the lowest fifth receiving 5 points and the highest fifth receiving 1 point. This index reflects overall plant-based food consumption regardless of healthfulness. The hPDI positively scored healthy plant-based foods, with the lowest fifth receiving 1 point and the highest fifth receiving 5 points and inversely scored both less healthy plant-based and animal-source food groups. This plant-based diet index reflects the greater intake of healthy plant-based foods and lower consumption of less healthy plant-based and animal-source foods. In contrast, the uPDI assigned higher scores to less healthy plant-based foods and lower scores to healthy plant foods and animal foods, with higher scores indicating greater consumption of less healthy plant-based foods. The total score for each index was obtained by summing the scores (ranging from 1 to 5) across all component food groups, resulting in a theoretical range from 18 to 90 points for each index.

The calculated scoring range in this study was 28–82 points for PDI, 29–82 points for hPDI, and 29–87 points for uPDI.

Food biodiversity using dietary species richness

Food biodiversity is conventionally partitioned into three components: richness, evenness, and disparity.[22] However, due to a lack of consensus on the measurement of evenness (the way different species are distributed in a sample in terms of relative abundance) and disparity (the degree of similarity between species in a diet, such as in their nutrition or taxonomy),[22] we used dietary species richness (DSR), which is the number of taxonomic species consumed by an individual within a specific time frame. Calculation of DSR as a metric of food biodiversity has been described by Hanley-Cook et al.[15] In brief, in EPIC, the calculation of DSR was based on the total number of distinct species consumed from individual foods, drinks, and mixed dishes (i.e., recipes were decomposed into ingredients) across the previous year.[15] Three scenarios of DSR were computed to assess the impact of consuming species in relatively small quantities. These scenarios consider either the total DSR or DSR excluding the 5 or 10% least-consumed species from each food group in EPIC, based on Daly et al.[22] For our analyses, total DSR was used. In a previous study, the exclusion of the least-consumed species did not alter the direction or magnitude of the relationship with mortality or cancer risk.[15]

The composite 3V score

We constructed a composite score, referred to as the 3V score, by standardising three key dietary exposures, namely UPF intake (% g/day), hPDI (18–90 points), and DSR (species/year). To account for differing units, each variable was standardised (z-score transformation), and the composite score was calculated by summing the standardised values of hPDI and DSR, then subtracting the standardised value of UPF intake, assuming each dimension has equal importance. Equal weighting was applied because no established evidence exists to support different weighting of the components.[17]

The cumulative 3V score

We developed a cumulative score using data-driven methods to maintain balanced groups across dietary patterns. The aim was to count the cumulative number of the 3V criteria met, after a binary classification based on highest (Q4 and Q5) vs lowest quintiles (<Q4) for each dietary pattern. Thresholds, corresponding to minimum values of Q4, were 13.96% g/day for UPFs, 73 for DSR, 58 for hPDI, and 55 for PDI. For DSR, hPDI, and PDI those surpassing these thresholds received a score of 1, and otherwise 0. For UPF, those surpassing the threshold received a score of 0, and otherwise 1. These scores were then summed to calculate adherence to the combination of 3V components giving a theoretical range for the cumulative 3V score of 0–3 criteria met. PDI was included only for sensitivity analysis; the main focus was on hPDI as the second pillar of the 3V diet.

Follow-up for colorectal cancer incidence

Colorectal cancer cases were identified until 31 December 2013, using cancer registries or during follow-up from a combination of sources including cancer and pathology centres, health insurance records, and active follow-up of study participants. The end of follow-up was determined as the latest date of follow-up for cancer incidence, death, or end of follow-up, whichever came first. Censoring dates for complete follow-up data from cancer registries were between December 2009 and December 2013.

Statistical analysis

Descriptive statistics

Descriptive data on anthropometric, socio-demographic, and lifestyle factors are presented as means (standard deviation (SD)) or frequency (percentages). Dietary variables, which were not normally distributed based on visual inspection of histograms and Q–Q plots, are summarised as medians (interquartile ranges (IQR)). Spearman's rank correlation (ρ) was used to assess the relationship between the aforementioned dietary scores and other known modifiable dietary factors previously associated with colorectal cancers, including fibre, dairy, calcium, red meat, processed meat and alcohol.[23]

Cox regression models

Dietary exposures, namely UPF intake, PDIs (PDI, hPDI, uPDI), DSR and 3V scores were analysed as both continuous and categorical variables. Absolute risk differences were calculated as the number of cases per 10,000 person-years in Q_5 and Q_1 of the 3V score, respectively.

For the categorical analysis, exposures were categorised into quintiles to assess potential dose–response relationships. Multivariable-adjusted Cox proportional hazard regression models were used to assess the associations between dietary exposures and colorectal cancer risk. Age was used as a primary time variable, and the end of follow-up was defined as the age at cancer diagnosis or at the last follow-up, whichever came first. The Wald test was used to evaluate trends within each Q.

For the regression models, we used a priori knowledge and a systematic covariate modelling strategy to explore the relationship between diet, lifestyle, and cancer risk, controlling for potential confounders. Specifically, Models 1 were stratified by sex (male, female), age at recruitment (in 1-year intervals), and centre; Models 2 were further adjusted for lifestyle-related

covariates, including alcohol intake at recruitment (g/day), education (none or primary school completed; secondary school; technical or professional school; longer education including university degree; not specified), smoking intensity (never; current, 1–15 cigarettes/day; current, 16–25 cigarettes/day; current, 26+ cigarettes/day; former, quit ≤10 years; former, quit 11–20 years; former, quit 20+ years; current, pipe/cigar/occasionally; unknown), physical activity (Cambridge index: inactive; moderately inactive; moderately active; active; missing), height (cm), body mass index (BMI) (kg/m^2), and energy intake (kcal/day) to account for known risk factors.

Single-value imputation was used to address missing data in covariates, controlling for potential confounding, such as physical activity and education level. Mode imputation was applied to baseline categorical covariates, with missingness of 3.7% for education level and 2% for physical activity, using the mode values "primary school completed" for education level and "moderately inactive" for physical activity. In the previous EPIC study, multiple imputation was used in sensitivity analyses, with results similar to those obtained using single-value imputation in the main analysis.[24] As a result, only single-value imputation was applied in this study. Missing data were assumed to be missing at random, meaning the probability of missingness depends on observed data, consistent with previous EPIC analyses.[24]

Model 2 was the main model used. Cox proportional hazards assumptions were confirmed using Schoenfeld residuals, based on follow-up time (years). Hazard ratios (HR) with 95% confidence intervals (CI) are presented for Model 2 throughout the manuscript unless otherwise stated.

Furthermore, we fitted mutually adjusted Cox proportional hazards models to assess the independent associations (i.e., Model 2, also including the two other dietary scores) between standardised UPF consumption, DSR or plant-based diet indices (i.e., one SD increments) with colorectal cancer incidence and its anatomical subsites.

We tested the effects of 3V diet criteria met on colorectal cancer risk by comparing combinations of criteria met as compared to none met. In addition, we examined the interaction terms between continuous dietary scores to further test for any interactive effects.

Sensitivity analysis

To address potential sex differences, we conducted sensitivity analyses using sex-specific quintiles of the 3V score and presented estimates by sex. Associations were examined in the overall population and by sex, with the same model specifications as the main analyses. We examined country-specific differences in the associations between UPF intake, DSR, hPDI, and composite 3V score and colorectal cancer risk by fitting stratified multivariable-adjusted Cox proportional hazards regression models. For the UK, we additionally stratified by centres (i.e., Cambridge, Oxford Health Conscious, and Oxford General).

To examine the robustness of the cumulative score definition, we additionally conducted a sensitivity analysis using median-based cut-offs instead of Q4 thresholds for each 3V component. We also repeated our main analyses using the PDI, rather than hPDI, as the second pillar of the 3V score for sensitivity analysis, which included recalculating the composite and cumulative 3V scores. To explore the potential impact of UPF intake within a plant-based diet on colorectal cancer risk, we conducted an additional analysis among participants with high adherence to the PDI, defined as those in the top two quartiles (Q4–Q5) of PDI. Within this subgroup, participants were further classified by UPF consumption into high (Q4–Q5) and low (Q1–Q3) intake groups. Cox proportional hazards regression models adjusted for confounders were used to compare colorectal cancer risk between high PDI participants with high vs low UPF intake. We also excluded participants who were censored during the first two years of follow-up to reduce the potential for reverse causation due to undiagnosed cancer at recruitment for the main analysis. Sensitivity analyses included a basic model adjusted for energy intake and a fully adjusted model excluding BMI to explore its potential mediating role. Finally, we constructed a weighted 3V score using the β coefficients (β = ln [HR]) from the overall colorectal cancer Cox models as weights for each component. The weights were rescaled relative to the smallest absolute β to preserve the relative contribution of each 3V component while keeping the overall score on a comparable scale to the original unweighted 3V score.

Statistical analyses were conducted using the R software (v 4.3.1). All tests were two-sided, and *P*-values <0.05 were considered statistically significant.

Role of funding source

The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Baseline characteristics

Over a mean (SD) follow-up period of 14.9 years (4), 6162 participants developed colorectal cancer (Table 1). Overall, 70.8% of participants were female, while the average age and BMI were 51 years (9.8) and 25.3 kg/m^2 (4.2) at the time of recruitment, respectively. At baseline, participants' [median (IQR)] energy intake was 1998.9 (1633.4–2437.7) kcal/day, UPF intake was 11.9% (7.5–17.9) by weight and 32.3% (20.0–43.1) by energy, the hPDI was 55.0 (51.0–59.0), and DSR was 68.0 (52.0–78.0) species/year. When comparing the Q5 of 3V

	Overall	Q1	Q2	Q3	Q4	Q5	*P*-test
Participants, n	450,111	90,023	90,022	90,022	90,022	90,022	
Number of incident colorectal cancer cases	6162 (1.4)	1338 (1.5)	1353 (1.5)	1290 (1.4)	1129 (1.3)	1052 (1.2)	<0.001
Age at recruitment, years	51.1 (9.8)	49.8 (10.3)	50.9 (9.9)	51.5 (9.7)	51.7 (9.5)	51.7 (9.2)	0.828
Female	318,686 (70.8)	59,188 (65.7)	59,218 (65.8)	62,343 (69.3)	66,751 (74.1)	71,186 (79.1)	<0.001
Country							<0.001
Denmark	55,014 (12.2)	9817 (10.9)	12,493 (13.9)	13,345 (14.8)	11,721 (13.0)	7638 (8.5)	
France	67,403 (15.0)	5700 (6.3)	10,998 (12.2)	13,564 (15.1)	16,302 (18.1)	20,839 (23.1)	
Germany	48,557 (10.8)	4504 (5.0)	6093 (6.8)	8343 (9.3)	11,114 (12.3)	18,503 (20.6)	
Italy	44,545 (9.9)	2732 (3.0)	5541 (6.2)	9125 (10.1)	13,130 (14.6)	14,017 (15.6)	
Norway	33,975 (7.5)	23,885 (26.5)	7883 (8.8)	1888 (2.1)	303 (0.3)	16 (0.0)	
Spain	39,989 (8.9)	7495 (8.3)	11,476 (12.7)	10,821 (12.0)	7205 (8.0)	2992 (3.3)	
Sweden	48,674 (10.8)	19,596 (21.8)	13,943 (15.5)	8447 (9.4)	4938 (5.5)	1750 (1.9)	
The Netherlands	36,538 (8.1)	7667 (8.5)	8745 (9.7)	8339 (9.3)	7201 (8.0)	4586 (5.1)	
United Kingdom	75,416 (16.8)	8627 (9.6)	12,850 (14.3)	16,150 (17.9)	18,108 (20.1)	19,681 (21.9)	
Tumour site[a]							
Colon	3841 (0.9)	845 (0.9)	825 (0.9)	783 (0.9)	708 (0.8)	680 (0.8)	<0.001
Proximal colon	1856 (0.4)	425 (0.5)	403 (0.5)	376 (0.4)	316 (0.4)	336 (0.4)	<0.001
Distal colon	1726 (0.4)	373 (0.4)	364 (0.4)	356 (0.4)	325 (0.4)	308 (0.3)	0.060
Other (overlapping or unspecified colon)	259 (0.1)	47 (0.1)	58 (0.1)	51 (0.1)	67 (0.1)	36 (0.0)	0.033
Rectum	2058 (0.5)	436 (0.5)	476 (0.5)	451 (0.5)	376 (0.4)	319 (0.4)	<0.001
Age at diagnosis, years	65.9 (8.6)	65.9 (9.0)	65.8 (8.7)	65.6 (8.2)	66.0 (8.6)	65.9 (8.6)	<0.001
Educational level							<0.001
None or primary school completed	126,615 (28.1)	29,665 (33.0)	28,061 (31.2)	26,485 (29.4)	23,451 (26.1)	18,953 (21.1)	
Technical/professional school	103,782 (23.1)	25,238 (28.0)	21,973 (24.4)	20,127 (22.4)	18,946 (21.0)	17,498 (19.4)	
Secondary school	93,910 (20.9)	18,261 (20.3)	17,694 (9.7)	18,019 (20.0)	19,291 (21.4)	20,645 (22.9)	
Longer education (incl. University degree)	108,931 (24.2)	14,767 (16.4)	19,149 (21.3)	21,713 (24.1)	24,425 (27.1)	28,877 (32.1)	
Missing	16,873 (3.7)	2092 (2.3)	3145 (3.5)	3678 (4.1)	3909 (4.3)	4049 (4.5)	
Smoking status and intensity							<0.001
Never	191,403 (42.5)	36,705 (40.8)	37,694 (41.9)	39,163 (43.5)	39,367 (43.7)	38,474 (42.7)	
Current, 1–15 cig/day	52,440 (11.7)	13,486 (15.0)	11,460 (12.7)	10,053 (11.2)	9224 (10.2)	8217 (9.1)	
Current, 16–25 cig/day	27,623 (6.1)	7512 (8.3)	6329 (7.0)	5347 (5.9)	4723 (5.2)	3712 (4.1)	
Current, 26+ cig/day	6559 (1.5)	1674 (1.9)	1460 (1.6)	1324 (1.5)	1147 (1.3)	954 (1.1)	
Former, quit ≤10 years	43,340 (9.6)	8882 (9.9)	8500 (9.4)	8480 (9.4)	8621 (9.6)	8857 (9.8)	
Former, quit 11–20 years	37,670 (8.4)	6334 (7.0)	6895 (7.7)	7424 (8.2)	7891 (8.8)	9126 (10.1)	
Former, quit 20+ years	36,845 (8.2)	6697 (7.4)	6767 (7.5)	7060 (7.8)	7727 (8.6)	8594 (9.5)	
Current, pipe/cigar/ occasionally	39,907 (8.9)	5010 (5.6)	7737 (8.6)	8474 (9.4)	8948 (9.9)	9738 (10.8)	
Unknown	14,324 (3.2)	3723 (4.1)	3180 (3.5)	2697 (3.0)	2374 (2.6)	2350 (2.6)	
Physical activity (Cambridge index)							<0.001
Inactive	88,032 (19.6)	15,405 (17.1)	17,722 (19.7)	18,517 (20.6)	18,788 (20.9)	17,600 (19.6)	
Moderately inactive	149,941 (33.3)	26,725 (29.7)	29,049 (32.3)	30,748 (34.2)	31,450 (34.9)	31,969 (35.5)	
Moderately active	120,199 (26.7)	28,495 (31.7)	24,213 (26.9)	22,242 (24.7)	22,183 (24.6)	23,066 (25.6)	
Active	83,115 (18.5)	16,317 (18.1)	16,814 (18.7)	16,922 (18.8)	16,424 (18.2)	16,638 (18.5)	
Missing	8824 (2.0)	3081 (3.4)	2224 (2.5)	1593 (1.8)	1177 (1.3)	749 (0.8)	
BMI, kg/m^2	25.3 (4.2)	25.5 (4.4)	25.5 (4.2)	25.4 (4.2)	25.1 (4.11)	24.8 (4.01)	<0.001
Height, cm	166.2 (9.0)	168.2 (8.8)	167.0 (9.1)	166.1 (9.0)	165.3 (8.7)	164.5 (8.2)	<0.001
Energy intake, kcal/day	1998.9 (1633.4–2437.7)	2116.9 (1704.7–2610.6)	2097.0 (1713.3–2545.0)	2040.0 (1677.7–2464.4)	1959.3 (1617.3–2360.8)	1817.1 (1501.4–2192.4)	<0.001

(Table 1 continues on next page)

	Overall	Q1	Q2	Q3	Q4	Q5	*P*-test
(Continued from previous page)							
Alcohol intake, g/day	5.5 (0.9–15.2)	2.9 (0.4–10.5)	5.3 (0.9–14.7)	6.2 (1.1–16.5)	6.7 (1.3–17.4)	6.8 (1.5–17.1)	<0.001
Fibre intake, g/day	21.9 (17.4–27.1)	20.3 (16.1–25.0)	21.3 (17.0–26.4)	22.0 (17.5–27.3)	22.5 (18.0–27.8)	23.5 (18.9–29.0)	<0.001
Calcium intake, mg/day	935.7 (712.5–1207.0)	905.1 (663.2–1199.1)	960.0 (723.4–1242.6)	965.1 (736.5–1241.0)	949.9 (732.3–1210.0)	902.0 (705.6–1144.1)	<0.001
Mediterranean diet score, 0–18 points	8.0 (6.0–11.0)	7.0 (5.0–9.0)	8.0 (5.0–10.0)	8.0 (6.0–10.0)	9.0 (7.0–11.0)	10.0 (8.0–12.0)	<0.001
Potatoes and other tubers, g/day	77.9 (43.5–127.5)	118.9 (63.5–158.0)	93.6 (55.0–141.7)	80.0 (46.0–128.0)	69.7 (36.3–112.9)	56.1 (27.1–87.3)	<0.001
Vegetables, g/day	167.3 (106.8–257.9)	120.2 (78.6–179.4)	152.2 (98.0–229.9)	173.9 (113.4–260.1)	192.0 (125.0–283.5)	219.2 (141.2–322.8)	<0.001
Legumes, g/day	5.4 (0.3–16.6)	0.0 (0.0–6.8)	3.3 (0.0–15.6)	6.2 (0.6–19.0)	6.6 (1.3–20.1)	8.5 (1.8–22.5)	<0.001
Fruit, nuts and seeds, g/day	193.1 (107.8–310.6)	127.6 (68.8–208.9)	167.0 (96.4–266.5)	199.2 (115.2–309.4)	231.7 (133.7–344.5)	268.9 (163.9–394.9)	<0.001
Cereal and cereal products, g/day	200.4 (143.1–271.7)	202.0 (146.9–263.9)	200.5 (144.5–269.4)	200.6 (143.3–273.9)	201.4 (143.0–278.7)	197.6 (137.9–275.2)	<0.001
Vegetable oils, g/day	2.5 (0.3–8.8)	0.7 (0.1–3.6)	1.6 (0.3–8.1)	2.9 (0.3–10.6)	4.1 (0.5–11.4)	4.9 (0.6–10.4)	<0.001
Fruit and vegetable juices, g/day	18.4 (0.5–94.3)	28.6 (0.0–107.1)	20.4 (0.1–100.1)	17.6 (0.9–94.8)	17.1 (1.2–93.7)	17.1 (1.2–78.6)	<0.001
Coffee, tea, herbal teas, g/day	540.0 (250.0–936.7)	450.0 (213.7–700.0)	540.0 (262.7–900.0)	580.2 (241.9–951.9)	600.0 (240.0–1000.0)	650.2 (292.9–1045.0)	<0.001
Dairy products, g/day	284.8 (166.1–454.9)	307.8 (174.3–497.7)	319.6 (188.8–497.2)	303.6 (180.3–466.5)	274.6 (163.1–427.6)	229.4 (131.4–365.0)	<0.001
Red meat, g/day	34.1 (15.6–62.8)	32.8 (17.2–63.6)	36.8 (17.0–68.6)	38.9 (16.9–67.7)	36.0 (16.0–62.8)	27.2 (10.1–50.0)	<0.001
Processed meat, g/day	25.9 (12.6–45.4)	38.0 (21.6–59.4)	30.3 (16.7–48.8)	25.4 (13.2–43.0)	21.1 (10.5–38.0)	16.8 (7.0–32.9)	<0.001
Poultry, g/day	14.7 (5.3–26.7)	14.0 (4.1–24.3)	15.2 (6.0–27.3)	15.7 (6.3–28.8)	15.5 (6.2–28.1)	12.6 (4.1–24.5)	<0.001
Fish and shellfish, g/day	28.6 (14.1–51.0)	38.3 (17.1–69.2)	31.5 (15.2–55.9)	28.5 (14.4–49.4)	26.4 (13.1–44.3)	22.6 (10.6–38.7)	<0.001
Egg and egg products, g/day	14.3 (7.0–23.8)	16.2 (8.1–27.5)	16.2 (7.4–27.4)	15.1 (7.1–24.6)	13.4 (6.7–22.8)	9.8 (4.9–19.2)	<0.001
Butter, g/day	0.3 (0.0–4.3)	0.1 (0.0–3.3)	0.2 (0.0–4.4)	0.4 (0.0–4.7)	0.6 (0.0–4.5)	0.6 (0.0–4.3)	<0.001
Proportion of Nova 1 in total diet, % g/day	73.6 (64.5–80.5)	65.5 (57.1–72.1)	72.4 (64.6–78.2)	74.9 (66.5–80.5)	76.6 (68.1–82.2)	79.3 (71.2–84.7)	<0.001
Proportion of Nova 2 in total diet, % g/day	0.9 (0.4–1.7)	0.8 (0.4–1.5)	0.9 (0.4–1.8)	0.9 (0.4–1.8)	0.9 (0.4–1.8)	0.9 (0.3–1.7)	<0.001
Proportion of Nova 3 in total diet, % g/day	10.6 (6.6–17.4)	9.5 (6.3–14.8)	10.6 (6.7–17.0)	11.0 (6.8–18.3)	11.2 (6.7–19.0)	11.1 (6.5–18.7)	<0.001
Proportion of Nova 4 (UPF) in total diet, % g/day	11.9 (7.5–17.9)	21.9 (15.8–29.0)	14.3 (9.7–19.2)	11.6 (7.8–15.9)	9.7 (6.5–13.4)	7.4 (4.9–10.6)	<0.001
Proportion of Nova 1 in total diet, % kcal/day	35.5 (28.7–42.6)	32.0 (26.1–38.2)	34.8 (28.5–41.5)	35.9 (29.4–42.8)	36.9 (30.1–44.0)	38.4 (30.6–46.2)	<0.001
Proportion of Nova 2 in total diet, % kcal/day	5.9 (2.3–11.4)	4.1 (1.9–8.4)	5.6 (2.2–10.9)	6.4 (2.3–11.9)	6.8 (2.5–12.4)	7.3 (2.7–12.7)	<0.001
Proportion of Nova 3 in total diet, % kcal/day	23.5 (15.1–32.7)	18.0 (11.8–26.2)	22.1 (14.5–30.4)	24.1 (16.0–32.8)	26.0 (17.3–34.9)	28.5 (18.5–37.8)	<0.001
Proportion of Nova 4 (UPF) in total diet, % kcal/day	32.3 (20.0–43.1)	43.0 (33.7–51.3)	35.1 (24.1–44.7)	31.2 (19.5–41.6)	27.9 (16.5–38.6)	23.3 (13.8–33.5)	<0.001
Plant-based diet index, 28–82 points	52.0 (49.0–56.0)	51.0 (47.0–55.0)	52.0 (48.0–55.0)	52.0 (49.0–56.0)	53.0 (49.0–56.0)	54.0 (51.0–58.0)	<0.001
Healthful plant-based diet index, 29–82 points	55.0 (51.0–59.0)	50.0 (46.0–54.0)	53.0 (50.0–56.0)	55.0 (52.0–58.0)	57.0 (55.0–60.0)	61.0 (59.0–64.0)	<0.001
Unhealthful plant-based diet index, 29–87 points	55.0 (51.0–60.0)	60.0 (56.0–64.0)	56.0 (52.0–60.0)	54.0 (51.0–59.0)	53.0 (50.0–57.0)	52.0 (48.0–56.0)	<0.001
DSR, number of species per year	68.0 (52.0–78.0)	50.0 (42.0–67.0)	63.0 (46.0–74.0)	69.0 (56.0–78.0)	72.0 (63.0–80.0)	76.0 (69.0–82.0)	<0.001

The composite 3V score was constructed by standardising three key dietary exposures: UPF intake (% g/day), hPDI (18–90 points), and DSR (species/year) to account for differences in their units of measurement. Each participant's 3V score was calculated by summing the standardised values of DSR and hPDI, and subtracting the standardised value of UPF intake, assuming equal importance for each dimension (i.e., 3V score = SD(DSR) + SD(hPDI)–SD(UPF)). Values are presented as mean (SD), median (IQR) for dietary variables, or frequency (%). *P*-values were calculated using ANOVA or Kruskal–Wallis tests for continuous variables and Chi-squared tests for categorical variables. [a]Values are presented as the number of incident colorectal cancer cases (frequencies) for each tumour site in the cohort. Among the total malignant colorectal cancer cases (6,162), colon, proximal colon, distal colon, and rectal account for 62.3%, 30.1%, 28.01%, and 33.4%, respectively. DSR, Dietary Species Richness; IQR, Interquartile Range; SD, Standard Deviation; UPF, Ultra-Processed Foods.

Table 1: **Participant characteristics of the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort (*n* = 450,111) by Quintiles (Qs) of the 3V score.**

score against the Q1, our findings indicate large differences in median (IQR) UPF intake (%g/d) (7.4% (4.9–10.6) vs 21.9% (15.8–29.0)), hPDI (61.0 (59.0–64.0) vs 50.0 (46.0–54.0)) and DSR (species/year) (76.0 (69.0–82.0) vs 50.0 (42.0–67.0)) (Table 1). Participants in Q5 of the 3V score were older (51.7 (9.2) vs 49.8 (10.3)), had a lower BMI (24.8 (4.0) vs 25.5 (4.4) kg/m^2), a higher education level (32.1% vs 16.4% with university education), and a greater proportion of never smokers (42.7% vs 40.8%).

hPDI was moderately negatively correlated with UPF intake (ρ = –0.32, P < 0.001) and very weakly with DSR (ρ = –0.09, P < 0.001), while DSR was very weakly positively associated with UPF intake (ρ = 0.18, P < 0.001) (Supplementary Figure S2). Weak correlations were found between the dietary scores and specific food group intakes or nutrients (i.e., fibre, dairy, calcium, red meat, processed meat and alcohol), with a moderate negative correlation reported between hPDI and red meat intake (ρ = –0.36, P < 0.001) (Supplementary Figure S2).

Associations between the composite 3V score and colorectal cancer risk

The 3V score showed significant inverse associations with colorectal cancer (HR per 1-SD increment: 0.96; 95% CI: 0.94–0.98), overall colon cancer (HR: 0.96; 95% CI: 0.93–0.98), and distal colon cancer risk (HR: 0.95; 95% CI: 0.92–0.99) (Fig. 1A and Supplementary Table S4). Absolute colorectal cancer risks among participants in the highest and lowest quintile of the 3V score were 8.59 and 10.37 cases/10,000 person-years, respectively (Supplementary Table S4). Comparing the highest and lowest quintile of the 3V score, we observed a significant inverse association with colorectal cancer ($HR_{Q5\ vs\ Q1}$: 0.84; 95% CI 0.76–0.94), and even stronger associations with colon (HR: 0.82; 95% CI: 0.72–0.93) and distal colon cancers (HR: 0.81; 95% CI: 0.67–0.99) (Fig. 1B and Supplementary Table S4). The analysis showed a dose–response relationship (P-trend <0.001).

Associations between 3V components and colorectal cancer risk

UPF intake

We observed a significant positive association between UPF intake and colorectal cancer risk, with consistent associations across models (Supplementary Table S5). On average, a 1 SD increment in UPF intake was associated with higher risk of colorectal cancer (HR: 1.06; 95% CI: 1.02–1.09), colon cancer (HR 1.07; 95% CI: 1.03–1.12), and distal colon cancer (HR 1.08; 95% CI: 1.02–1.15) (Fig. 1A and Supplementary Table S5). Comparing the highest and lowest quintile of UPF intake, we found significant positive associations with colorectal cancer ($HR_{Q5\ vs\ Q1}$: 1.15; 95% CI: 1.04–1.27), colon cancer (HR: 1.20; 95% CI: 1.05–1.36), and distal colon cancer (HR: 1.22; 95% CI: 1.01–1.47). No significant associations were observed for rectal cancer (Fig. 1B and Supplementary Table S5). The analysis showed a dose–response relationship (P-trend <0.001) (Fig. 1B and Supplementary Table S5).

PDIs

Higher hPDI was associated with lower risk of colorectal cancer (HR per 1-SD increment: 0.95; 95% CI: 0.91–0.98) (Fig. 1A and Supplementary Table S5). The associations were slightly stronger for distal colon cancer (HR: 0.93; 95% CI: 0.87–0.99) compared to other sites. Comparing the highest and lowest quintile of hPDI, we observed significant inverse associations with colorectal ($HR_{Q5\ vs\ Q1}$: 0.82; 95% CI: 0.74–0.91), colon (HR: 0.82; 95% CI: 0.73–0.94), distal colon (HR: 0.74; 95% CI: 0.61–0.90) and rectal cancers (HR: 0.82; 95% CI: 0.69–0.98). The analysis showed a dose–response relationship (P-trend <0.001) (Fig. 1B and Supplementary Table S5).

We did not observe significant associations between either continuous or categorical PDI and colorectal cancer risk (Supplementary Table S6). However, for uPDI, positive associations were observed with colorectal (HR per 1-SD increment: 1.06; 95% CI: 1.03–1.10), colon (HR: 1.09; 95% CI: 1.04–1.13), distal colon (HR: 1.09; 95% CI: 1.02–1.17), and proximal colon cancers (HR: 1.07; 95% CI: 1.01–1.14) (Supplementary Figure S3A and Supplementary Table S6). Comparing the highest and lowest quintile of uPDI, we found significant positive associations with colorectal cancer ($HR_{Q5\ vs\ Q1}$: 1.18; 95% CI: 1.06–1.31), colon cancer (HR: 1.26; 95% CI: 1.10–1.43), and proximal colon cancer (HR: 1.25; 95% CI: 1.04–1.51) (Supplementary Figure S3B and Supplementary Table S6).

DSR

DSR was inversely associated with colorectal cancer (HR per 1-SD increment: 0.95; 95% CI: 0.90–0.99) (Fig. 1A and Supplementary Table S5), while weak associations were observed with colon and proximal colon cancers in crude models (Supplementary Table S5). Comparing the highest and lowest quintile of DSR, we found significant inverse association with colorectal cancer ($HR_{Q5\ vs\ Q1}$: 0.84; 95% CI: 0.72–0.99) (Fig. 1B and Supplementary Table S5). For proximal colon cancer, the association was stronger (HR: 0.70; 95% CI: 0.53–0.93)). The analysis showed a dose–response relationship (P-trend <0.001) (Fig. 1B and Supplementary Table S5).

Mutually adjusted associations

Mutually adjusted associations were statistically significant and positive, but attenuated, for UPF intake and colorectal cancer (HR per 1-SD increment: 1.05; 95% CI: 1.01–1.08) and colon cancer (HR: 1.06; 95% CI: 1.02–1.11) (Fig. 2A and Supplementary Table S7). For DSR, we observed a slightly strengthened inverse

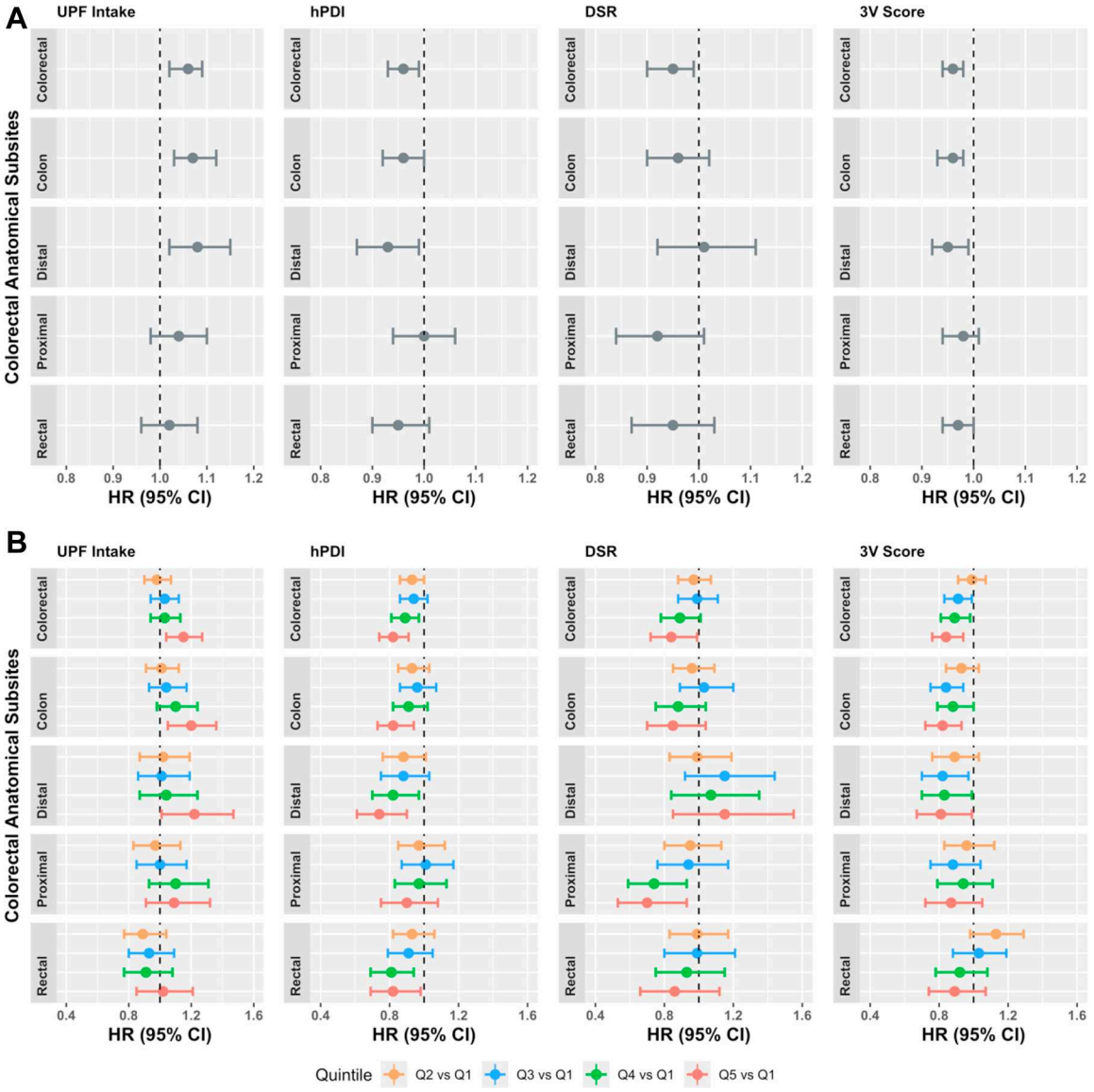

Fig. 1: **Forest plot of the associations between dietary scores—UPF intake, DSR, hPDI, and the 3V score (calculated as SD(DSR) + SD (hPDI)–SD(UPF))—and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort (*n* = 450,111).** Panel A indicates the hazard ratios (HR) for a 1 SD increase in UPF intake, DSR, hPDI and the 3V score, while panel B indicates the HR comparing each dietary score quintile (Q2, Q3, Q4, Q5) to the lowest quintile (Q1) for UPF intake, DSR, hPDI, and the 3V score. The composite 3V score was constructed by standardising three key dietary exposures: UPF intake (% g/day), hPDI (18–90 points), and DSR (species/year) to account for differences in their units of measurement. Each participant's 3V score was calculated by summing the standardised values of DSR and hPDI, and subtracting the standardised value of UPF intake, assuming equal importance for each dimension (i.e., 3V score = SD (DSR) + SD (hPDI)–SD (UPF)). Regression models were separately fitted for each dietary score, and the same model adjustments were applied for each score. The models were stratified for centre, sex (male, female), and age at recruitment (1-year intervals, timescale) and adjusted for education (none or primary school completed; secondary school; technical or professional school; longer education including university degree; not specified), smoking status and intensity of smoking (never; current, 1–15 cigarettes/day; current, 16–25 cigarettes/day; current, 26+ cigarettes/day; former, quit ≤10 years; former, quit 11–20 years; former, quit 20+ years; current, pipe/cigar/occasionally; unknown), physical activity (Cambridge index: inactive; moderately inactive; moderately active; active; missing), height (cm), body mass index (kg/m^2), alcohol intake at recruitment (g/day), and energy intake (kcal/day). N = 450,111, with 6162 cases of colorectal cancer, 3841 cases of colon cancer, 1726 cases of distal colon cancer, 1856 cases of proximal colon cancer, and 2058 cases of rectal cancer. CI, Confidence Interval; DSR, Dietary Species Richness; hPDI, healthful Plant-Based Diet Index; SD, Standard Deviation; UPF, Ultra-Processed Foods.

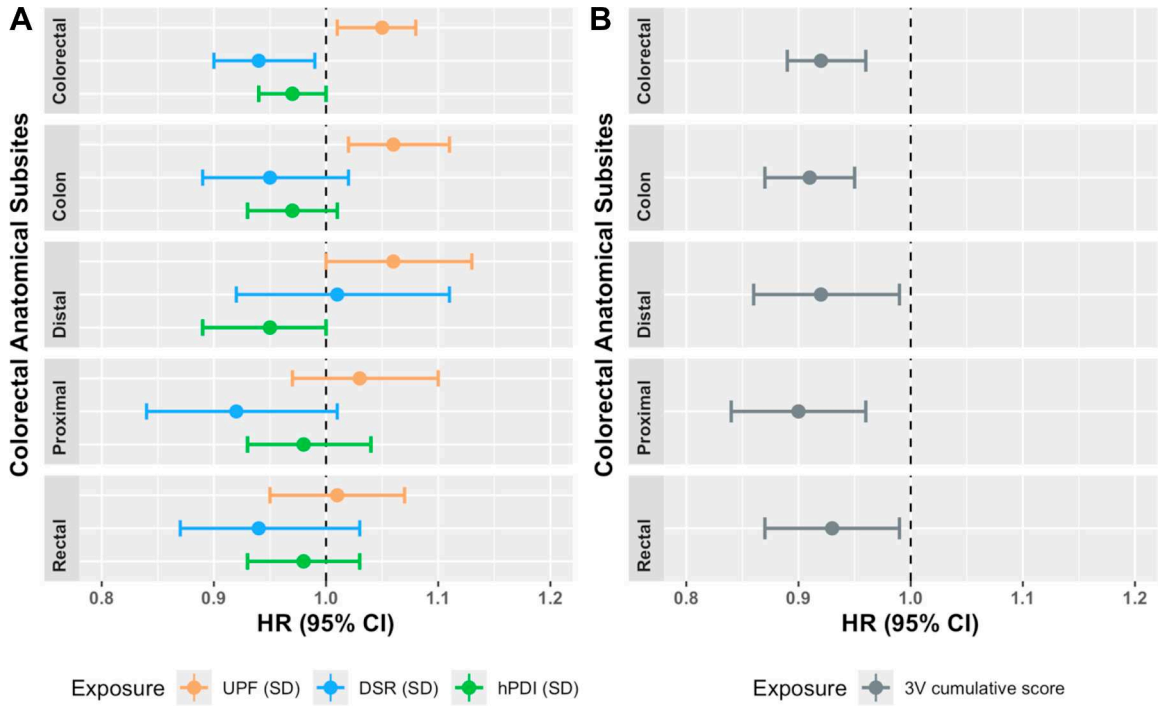

Fig. 2: **Forest Plot of the associations between UPF intake, DSR, hPDI and cumulative 3V score (calculated as DSR (0 or 1), hPDI (0 or 1), and UPF (1 or 0)) and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort (*n* = 450,111).** Panel A indicates the associations of dietary scores when mutually adjusted. Panel B indicates the association of the cumulative 3V score (HR per criterion met). The models were stratified for centre, sex (male, female), and age at recruitment (1-year intervals, timescale) and adjusted for education (none or primary school completed; secondary school; technical or professional school; longer education including university degree; not specified), smoking status and intensity of smoking (never; current, 1–15 cigarettes/day; current, 16–25 cigarettes/day; current, 26+ cigarettes/day; former, quit ≤10 years; former, quit 11–20 years; former, quit 20+ years; current, pipe/cigar/occasionally; unknown), physical activity (Cambridge index: inactive; moderately inactive; moderately active; active; missing), height (cm), body mass index (kg/m^2), alcohol intake at recruitment (g/day), and energy intake (kcal/day). Thresholds for the cumulative 3V score were set at the minimum values of the fourth quintile (Q4) for each dietary component: 13.96% g/day for UPF (scored as 1 for lower consumption and 0 for higher), 58 points for hPDI, and 73 species/year for DSR, with participants in Q4 and Q5 receiving a score of 1 for both. These scores were then summed to calculate adherence to the combination of 3V components, giving a theoretical range for the cumulative 3V score of 0–3 criteria met (cumulative 3V score = DSR (0 or 1) + hPDI (0 or 1) + UPF (0 or 1)). *n* = 450,111, with 6162 cases of colorectal cancer, 3841 cases of colon cancer, 1726 cases of distal colon cancer, 1856 cases of proximal colon cancer, and 2058 cases of rectal cancer. CI, Confidence Interval; DSR, Dietary Species Richness; HR, Hazard Ratio; hPDI, healthful Plant-Based Diet Index; SD, Standard Deviation; UPF, Ultra-Processed Foods.

association with colorectal cancer (HR: 0.94; 95% CI: 0.90–0.99) (Fig. 2A and Supplementary Table S7). For hPDI, we found an attenuated, non-significant inverse association with colorectal, colon, and distal colon cancer risk (Fig. 2A and Supplementary Table S7). Furthermore, when we tested the interactions among these dietary scores, we did not find statistically significant relationships with colorectal cancer (Supplementary Figure S4).

Associations between the cumulative 3V score and colorectal cancer risk

The cumulative 3V score showed significant inverse associations with colorectal cancer (HR per criteria met: 0.92; 95% CI: 0.89–0.96), overall colon cancer (HR: 0.91; 95% CI: 0.87–0.95), proximal colon cancer (HR: 0.90; 95% CI: 0.84–0.96), distal colon cancer (HR: 0.92; 95% CI: 0.86–0.99) and rectal cancer (HR: 0.93; 95% CI: 0.87–0.99) (Fig. 2B and Supplementary Table S8).

Adherence to combination of 3V components and colorectal cancer risk

Meeting the low UPF intake threshold was inversely associated with colorectal cancer risk, as compared to meeting none of the 3V diet thresholds (HR: 0.91; 95% CI: 0.83–0.99) (Fig. 3 and Supplementary Table S9). Meeting the high hPDI (HR: 0.93; 95% CI: 0.80–1.08) or DSR (HR: 0.87; 95% CI: 0.75–1.01) thresholds were not significantly associated with colorectal cancer risk (Fig. 3). Combined adherence to low UPF and high

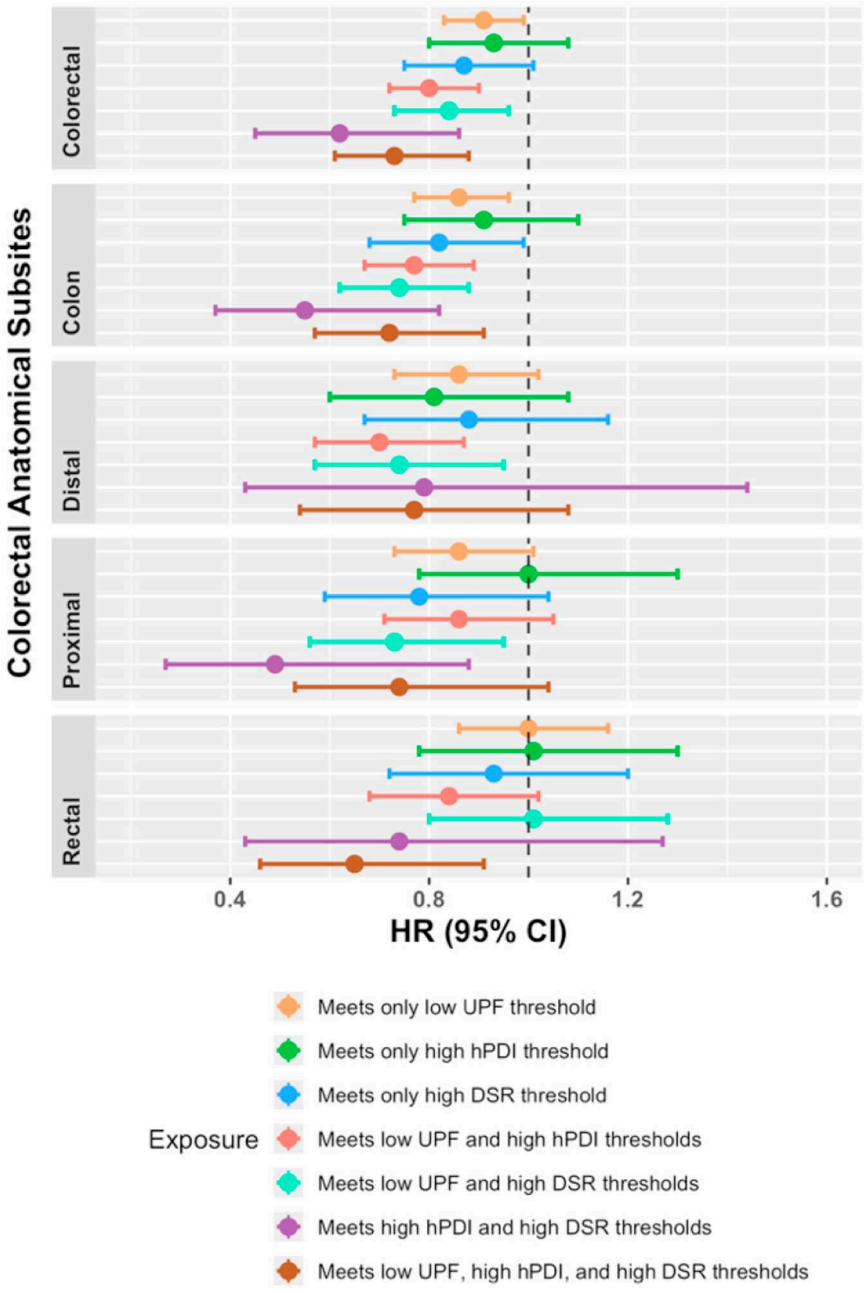

***Fig. 3:* Associations between UPF, DSR, and hPDI thresholds and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort (*n* = 450,111)**. All the associations were compared to the scenario where none of the recommendations were met. Thresholds for the cumulative 3V score were set at the minimum values of the fourth quintile (Q4) for each dietary component: 13.96% g/day for UPF (scored as 1 for lower consumption and 0 for higher), 58 points for hPDI, and 73 species/year for DSR, with participants in Q4 and Q5 receiving a score of 1 for both. The models were stratified for centre, sex (male, female), and age at recruitment (1-year intervals, timescale) and adjusted for education (none or primary school completed; secondary school; technical or professional school; longer education including university degree; not specified), smoking status and intensity of smoking (never; current, 1–15 cigarettes/day; current, 16–25 cigarettes/day; current, 26+ cigarettes/day; former, quit ≤10 years; former, quit 11–20 years; former, quit 20+ years; current, pipe/cigar/occasionally; unknown), physical activity (Cambridge index: Inactive; moderately inactive; moderately active; active; missing), height (cm), body mass index (kg/m^2), alcohol intake at recruitment (g/day), and energy intake (kcal/day). The numbers of events/total participants by cancer site and threshold, among those who meet the threshold for each criterion, are as follows: For colorectal cancer, low UPF (2593/168,947), high hPDI (1353/103,157), high DSR (1933/136,185), low UPF and high hPDI (2059/157,897), low UPF and high DSR (1867/130,325), high hPDI and high DSR (1187/90,121), and low UPF, high hPDI, and high DSR (1500/111,739). For colon, low UPF (1623/

hPDI or low UPF and high DSR (HR: 0.80; 95% CI: 0.72–0.90 and HR: 0.84; 95% CI: 0.73–0.96, respectively) as well as adherence to both high hPDI and high DSR showed inverse associations with colorectal cancer risk (HR: 0.62; 95% CI: 0.45–0.86) (Fig. 3). Adherence to all three criteria—low UPF, high hPDI, and high DSR—showed inverse associations with colorectal cancer (HR: 0.73, 95% CI: 0.61–0.88) (Fig. 3). Similar directions and magnitudes of associations were observed across the anatomical sites (Fig. 3).

Sensitivity analysis

Higher scores were inversely associated with colorectal cancer and its subtypes in both men and women, with statistically significant trends observed for colorectal and colon cancer outcomes (Supplementary Tables S10 and S11). Across countries, UPF intake showed a consistent positive association with colorectal cancer risk, appearing more pronounced in Denmark. Inverse associations for greater DSR were consistent, with the largest relationship found in the UK. Moreover, higher hPDI showed consistent inverse associations with colorectal cancer, appearing more pronounced in Spain and the Netherlands. Similarly, the 3V score's inverse association was most pronounced in Denmark and Spain. Comparable directions and magnitudes of associations were observed across countries and centres (Supplementary Figure S5 and Supplementary Table S12). Similar results were observed when the participants who were censored in the first two years of follow-up were excluded (Supplementary Table S13). The weighted 3V score produced effect estimates nearly identical to the unweighted score (HR for overall colorectal cancer: 0.96, 95% CIs 0.95–0.98), confirming the robustness of the main findings (Supplementary Table S14).

PDI was not significantly associated with colorectal cancer risk (Supplementary Figure S6). In the sensitivity analysis using median-based cut-offs, inverse associations with colorectal and rectal cancer risk remained generally consistent, while associations for colon subtypes were weaker and less consistent (Supplementary Table S15). For the 3V score, when including PDI rather than hPDI inverse associations with colorectal cancer (HR per 1-SD increment: 0.97; 95% CI: 0.95–0.99), and weaker associations with colon (HR: 0.96; 95% CI: 0.94–0.99) and distal colon cancer risk (HR: 0.96; 95% CI: 0.92–1.00) were observed (Supplementary Figure S7 and Supplementary Table S16). Furthermore, the significant associations identified in the 3V score quintile analysis were no longer significant when PDI rather than hPDI was included in the analysis. Among participants with high adherence to the PDI (Q4–Q5), there was no statistically significant difference in colorectal cancer risk between those with high vs low UPF intake (HR: 1.05, 95% CI: 0.95–1.18). This suggests that within a high plant-based diet, elevated UPF consumption does not substantially alter colorectal cancer risk in this cohort (Supplementary Table S17). Similar results were observed in sensitivity analyses comparing the energy-adjusted basic model and the model without BMI to the crude and fully adjusted models (Supplementary Table S18).

Moreover, for the cumulative 3V score, most associations weakened using PDI rather than hPDI, with statistical significance no longer observed for distal colon and rectal cancer, while only the association with proximal colon cancer became slightly stronger (Supplementary Figure S8 and Supplementary Table S19).

Discussion

This is the first study to investigate the associations between a 3V diet—characterised by low UPF, healthy plant-based and biodiverse—and colorectal cancer risk in a large prospective cohort. Higher 3V scores demonstrated consistent and robust inverse associations with colorectal cancer risk and its anatomical subsites. Furthermore, our findings indicated significant associations between three complementary dietary scores, namely UPF intake, hPDI, and DSR, and colorectal cancer risk. Specifically, UPF intake was positively associated with colorectal and colon cancers when mutually adjusted for the other 3V components. Greater DSR was inversely associated with colorectal cancer risk when mutually adjusted for the other 3V components. hPDI was inversely associated with colorectal, colon and distal colon cancer risk, whereas these associations were not significant when mutually adjusted for the other 3V components. Adherence to

167,919), high hPDI (855/102,643), high DSR (1207/135,426), low UPF and high hPDI (1328/157,105), low UPF and high DSR (1139/129,544), high hPDI and high DSR (751/89,664), and low UPF, high hPDI, and high DSR (958/111,158). For distal colon, low UPF (734/167,030), high hPDI (368/102,156), high DSR (524/134,743), low UPF and high hPDI (599/156,376), low UPF and high DSR (515/128,920), high hPDI and high DSR (337/89,250), and low UPF, high hPDI, and high DSR (437/110,637). For proximal colon, low UPF (791/167,087), high hPDI (433/102,221), high DSR (604/134,823), low UPF and high hPDI (654/156,431), low UPF and high DSR (546/128,951), high hPDI and high DSR (372/89,285), and low UPF, high hPDI, and high DSR (466/110,666). For rectal cancer, low UPF (884/167,180), high hPDI (453/102,241), high DSR (644/134,863), low UPF and high hPDI (660/156,437), low UPF and high DSR (640/129,045), high hPDI and high DSR (393/89,306), and low UPF, high hPDI, and high DSR (478/110,678). CI, Confidence Interval; DSR, Dietary Species Richness; HR, Hazard Ratio; hPDI, healthful Plant-Based Diet Index; SD, Standard Deviation; UPF, Ultra-Processed Foods.

low UPF, high hPDI, and high DSR was inversely associated with colorectal, colon, and rectal cancer risk compared to those adhering to none.

These results align with prior observational studies, which have reported positive associations between higher UPF consumption and colorectal cancer, including findings from the EPIC cohort,[11] inverse associations for plant-based diets,[12–14] particularly those assessed using the hPDI, and associations between greater food biodiversity and lower gastrointestinal cancer risk in Europe as reported in the EPIC cohort.[15,16] Mutual adjustment between continuous UPF intake, PDIs, and DSR scores led to only weak attenuation of the relationships, indicating that the three pillars of the 3V score are largely independent. In this mutually adjusted model, UPF intake appeared a stronger driver of the positive association with colorectal and colon cancer risk, while the other components of the 3V framework, including the plant-based diet and food biodiversity, showed weaker or non-significant associations, particularly for colon cancer. The association observed for hPDI and colorectal cancer became non-significant after adjusting for UPF and DSR. This attenuation might be primarily influenced by UPF intake, as suggested by the moderate negative correlation between UPF and hPDI.

Adherence to each of the three pillars of the cumulative 3V score, as compared to none, was associated with a 27% reduction for colorectal cancer, 28% for colon and 35% for rectal cancer. We observed an additive reduction in risk for additional criteria met and no evidence for synergistic interaction between the 3V pillars. Among those only meeting one criterion, we observed a significant effect for meeting the UPF criterion, but not for those meeting the DSR or hPDI criteria, which may point to a relatively greater importance of UPF for colorectal cancer risk. However, the confidence intervals for effects of meeting each criterion and combinations thereof were generally overlapping, so observed differences in effect sizes should be treated with caution.

hPDI and the 3V score showed the strongest inverse associations with distal colon cancer. This might be related to the fact that different anatomical subsites of the colon-rectum may have varying susceptibilities to (processed) meat intake and potentially other dietary risk factors.[25] Etemadi et al.[25] indicated that the distal colon and rectum are more susceptible to the influence of red meat and haem iron intake, suggesting that the potential causes of colorectal cancer vary across anatomical subsites. Additionally, another study based on EPIC data showed differences in the impact of major risk factors (physical activity, anthropometric measurements, and smoking) on colorectal cancer across anatomical subsites.[26] Our study confirms this heterogeneity, particularly in the distal colon region, which showed significant negative associations with 3V scores. This suggests that the distal colon is more sensitive to dietary factors, especially in the context of a comprehensive dietary profile. However, the DSR association was less pronounced in the distal colon, potentially due to a weak positive correlation with red meat intake, which may diminish the protective effects of food biodiversity on this site.

Possible mechanisms for the potential protective effects of the 3V diet may include higher fibre intake and greater diversity of phytochemicals, which in turn could enhance the diversity of healthy gut microbiome; intake of low glycaemic index foods that might reduce inflammation and insulin resistance; and more adequate intakes of micronutrient that may lower oxidative stress.[27] Moreover, reducing ultra-processed food consumption may reduce exposure to xenobiotics, such as neo-formed contaminants (e.g. trans fats), colour additives (e.g. titanium dioxide), artificial sweeteners, and emulsifiers, consequently lowering colorectal cancer risk.[28] These possible mechanisms warrant further exploration.

Our large-scale investigation provided a comprehensive examination of dietary risk factors for colorectal cancer at various anatomical subsites and benefited from considerable statistical power, long-term follow-up, and substantial dietary heterogeneity. Other key strengths include the prospective design of the EPIC cohort, along with its pan-European nature, enhancing generalizability. Furthermore, cancer cases were identified through cancer registries and other robust methods, which ensured data quality. The use of validated country-specific dietary assessment methods enhances the reliability of dietary intake data across the diverse EPIC centres, and the availability of various lifestyle and socioeconomic variables allows for comprehensive adjustments, further strengthening the findings. Additionally, the Nova coding[19] and DSR coding[15] used in the EPIC cohort were developed by a team of international experts.

Nonetheless, our study is also subject to limitations. First, caution is warranted when generalising these findings to the broader European population or to other populations and ethnic groups, as the cohort comprised middle-aged volunteers from nine European countries enrolled in a long-term study on nutrition and health, who likely exhibit more health-conscious behaviours than the general population. EPIC participants were recruited from 10 Western European countries, and as no information on ethnicity was collected, caution is warranted in extrapolating the results to other populations worldwide or ethnic groups in Western Europe. Furthermore, only baseline data, collected in the nineties, were used for these analyses, while diets of the participants may have changed over time. For instance, the Nova coding was applied to dietary data collected over 20 years ago,[18] which may not reflect current high levels of UPF consumption.

Additionally, the study participants consisted of approximately 70% females, which may limit the generalizability of the findings to males. While this study indicates that UPFs contributed an average of 32% to daily energy intake based on historical data, their contribution could reach 60% in the UK.[29] The increasing trend in UPF consumption over time suggests that the impact of UPFs on cancer risk may be greater in contemporary society, potentially affecting the observed associations with the 3V score. Previous research has reported that over 30% of energy intake in vegetarian and vegan diets comes from UPFs,[30] which suggests a potential overlap between plant-based diets and UPFs. Some foods commonly consumed in plant-based diets—such as breakfast cereals—are also classified as ultra-processed, which may complicate the interpretation of the 3V score. However, in our additional analysis restricted to participants with high adherence to a plant-based diet, colorectal cancer risk did not differ by level of UPF intake, suggesting this overlap did not substantially impact the observed associations. UPFs capture nutrient composition degradation due to food processing itself, empty calories, artificial ingredients, foods with high glycaemic potential, and less satiating foods (high in fat and sugar, low in protein and dietary fibre). While many epidemiological studies associate high UPF intake with adverse health outcomes, the Nova classification has been criticised for including heterogeneous concepts in the UPF group (due to various nutritional compositions such as and variability between brands).[31] For instance, a study suggests that not all UPFs carry equal health risks (e.g. packaged long-shelf breads, flavoured yoghurts).[32] Despite these limitations, Nova remains the most widely used framework for capturing food processing in large-scale studies. We therefore used it alongside hPDI and DSR to capture complementary aspects of diet relevant to both health and sustainability, beyond what nutrient-based indices alone can assess.

In addition, DSR is likely underestimated as many broad FFQ items could not be decomposed into taxonomic species, although most were accounted for. The DSR metric reflects food diversity in terms of Linnaean taxonomy, and other diversity measures based on functional food groups, such as FGDS,[33] may show different associations. Furthermore, certain broad FFQ items could not be classified as exclusively plant- or animal-based (e.g. soup can be plant- or animal-based). Additionally, reliance on predominantly self-reported dietary data from a single baseline FFQ introduces the risk of exposure misclassification; however, the implementation of validated dietary assessment methods aims to mitigate this limitation. All components of the composite 3V score rely on self-reported dietary data, which inherently include measurement error. Without an established gold standard for each component, it is not possible to quantify or model their individual uncertainties. This limitation may affect the precision of the composite score and should be considered when interpreting the study's findings. Given EPIC's multicentre structure, dietary intake was assessed using different validated tools across centres, including self- or interviewer-administered FFQs, and in Malmö, a 7-day dietary record combined with a non-quantitative FFQ.[19] Despite rigorous harmonisation procedures, variation in assessment methods may have introduced exposure misclassification. However, as this variation is unlikely to be related to disease status, any resulting bias would likely attenuate true associations rather than create spurious ones. The FFQ, employed as the dietary assessment instrument in most countries, provides a closed list of food items. However, we cannot exclude the possibility that participants consumed other items not captured by the instrument. All these sources of misclassification are likely to lead to underestimation of the associations. Although we adjusted for known confounders, residual confounding from unmeasured or imprecisely measured factors cannot be excluded and may have influenced the observed associations. A further limitation is that the study is based on follow-up data available only through 2013, with more recent data not yet accessible to researchers, potentially leading to underestimation of associations by exclusion of more recent cases.

In summary, this study highlights the multifactorial relationship between diet and colorectal cancer risk, emphasizing that higher 3V scores—Veritable, Vegetal, and Varied diets—are significantly associated with lower colorectal cancer incidence and its anatomical subsites. Furthermore, we observed a cumulative reduction in risk for each 3V criterion met. The 3V score, by encouraging the consumption of non-ultra-processed, biodiverse and nutritionally healthy plant-based foods, aligns with current official recommendations in several countries such as Brazil and France. The present study underlines the importance of these largely independent dietary dimensions in colorectal cancer aetiology and, more generally, for preventive nutrition. The 3V score highlights key pillars of the diet-disease relationship and reflects dietary components that are relevant to broader sustainability issues (e.g. biodiversity loss). Consequently, public health policies and programs should facilitate reductions in UPF consumption, promoting nutritionally healthy plant-based diets, and enhancing food biodiversity.

Contributors

Conceptualization: Emine Koc Cakmak, Oliver J K Robinson, Paolo Vineis, Inge Huybrechts, Giles T. Hanley-Cook, Marc J Gunter. Data curation: Inge Huybrechts, Geneviève Nicolas, Carine Biessy. Formal analysis: Emine Koc Cakmak. Funding acquisition: Inge Huybrechts, Paolo Vineis, Giles T. Hanley-Cook, Carl Lachat, Mathilde Touvier.

Investigation: Emine Koc Cakmak, Inge Huybrechts. Methodology: Emine Koc Cakmak, Oliver J K Robinson, Paolo Vineis, Inge Huybrechts. Supervision: Oliver J K Robinson, Paolo Vineis, Inge

Huybrechts. Visualisation: Emine Koc Cakmak. Writing original draft: Emine Koc Cakmak. Writing paper draft: Emine Koc Cakmak, Oliver J K Robinson, Paolo Vineis, Inge Huybrechts, Marc J Gunter, Aline Al Nahas, Bernadette Chimera, Giles Hanley-Cook, Jeroen Berden, Anthony Fardet, Edmond Rock, Carine Biessy, Geneviève Nicolas, Nathalie Kliemann, Fernanda Rauber, Renata Bertazzi Levy, Lorenzo Mangone, Mathilde Touvier, Bernard Srour, Emmanuelle Kesse-Guyot, Carl Lachat, Guri Skeie, Elisabete Weiderpass, Franziska Jannasch, Christina C. Dahm, Daniel Borch Ibsen, Christina Dahl, Cecilie Kyrø, Mariem Hajji-Louati, Chloé Marques, Gianluca Severi, Verena Katzke, Rudolf Kaaks, Matthias B. Schulze, Saverio Caini, Sabina Sieri, Maria Santucci De Magistris, Rosario Tumino, Carlotta Sacerdote, Raúl Zamora-Ros, Maria-José Sánchez, Ana Jimenez-Zabala, Jesús-Humberto Gómez, Marcela Guevara, Elio Riboli. Emine Koc Cakmak and Inge Huybrechts have verified the underlying data in the study. All authors contributed to the editing and reviewing of the manuscript, and all have read and approved the final manuscript version.

Data sharing statement

Data described in the manuscript will be made available upon request pending approval by IARC and the EPIC centres. For application details and to request data described in the manuscript, please follow the instructions here: https://epic.iarc.fr/access/.

Declaration of interests

MT received public funding from the European Research Council, from the French National Institute for Health and Medical Research Inserm/France 2030 to her institution.

Acknowledgements

The authors thank all EPIC centres for provision of participant data. The coordination of EPIC-Europe is financially supported by International Agency for Research on Cancer (IARC) and also by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London (United Kingdom) which has additional infrastructure support provided by the NIHR Imperial Biomedical Research Centre (BRC). The national cohorts are supported by: Danish Cancer Society (Denmark); Ligue Nationale Contre le Cancer, Institut Gustave-Roussy, Mutuelle Générale de l'Education Nationale (MGEN), Institut National de la Santé et de la Recherche Médicale (INSERM), French National Research Agency (ANR, reference ANR-10-COHO-0006), French Ministry for Higher Education (subsidy 2102918823, 2103236497, and 2103586016) (France); German Cancer Aid, German Cancer Research Center (DKFZ), German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Federal Ministry of Education and Research (BMBF) (Germany); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy, Italian Ministry of Health, Italian Ministry of University and Research (MUR), Compagnia di San Paolo (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), the Netherlands Organisation for Health Research and Development (ZonMW), World Cancer Research Fund (WCRF), (The Netherlands); UiT The Arctic University of Norway; Health Research Fund (FIS)–Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology–ICO (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (C864/A14136 to EPIC-Norfolk; C8221/A29017 to EPIC-Oxford), Medical Research Council (MR/N003284/1, MC-UU_12015/1 and MC_UU_00006/1 to EPIC-Norfolk (DOI 10.22025/2019.10.105.00004); MR/Y013662/1 to EPIC-Oxford) (United Kingdom). Previous support has come from "Europe against Cancer" Programme of the European Commission (DG SANCO). We also thank the National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands, for their contribution and ongoing support to the EPIC Study. This work was supported by Cancer Research UK (C33493/A29678), World Cancer Research Fund International (IIG_FULL_2020_033) and the Institut National du Cancer (INCa number 2021–138). Funding for grant number IIG_FULL_2020_034 was obtained from Wereld Kanker Onderzoek Fonds (WKOF), as part of the World Cancer Research Fund International grant programme. EKC was supported by the Ministry of National Education, Türkiye, through the YLSY International Graduate Education Scholarship Programme. Researchers were independent from the funders. MT received public funding from the European Research Council, from the French National Institute for Health and Medical Research Inserm/France 2030 to her institution. 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.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2025.103662.

References

1 Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209–249.
2 Phipps AI, Lindor NM, Jenkins MA, et al. Colon and rectal cancer survival by tumor location and microsatellite instability: the Colon cancer family registry. *Dis Colon Rectum*. 2013;56:937.
3 Kocarnik JM, Shiovitz S, Phipps AI. Molecular phenotypes of colorectal cancer and potential clinical applications. *Gastroenterology*. 2015;3:269–276.
4 Siegel R, DeSantis C, Jemal A. Colorectal cancer statistics, 2014. *CA Cancer J Clin*. 2014;64:104–117.
5 Brenner H, Chen C. The colorectal cancer epidemic: challenges and opportunities for primary, secondary and tertiary prevention. *Br J Cancer*. 2018;119:785–792.
6 Veettil SK, Wong TY, Loo YS, et al. Role of diet in colorectal cancer incidence: umbrella review of meta-analyses of prospective observational studies. *JAMA Netw Open*. 2021;4:e2037341.
7 Wang L, Du M, Wang K, et al. Association of ultra-processed food consumption with colorectal cancer risk among men and women: results from three prospective US cohort studies. *BMJ*. 2022;378. https://doi.org/10.1136/bmj-2021-068921.
8 El Kinany K, Huybrechts I, Hatime Z, et al. Food processing groups and colorectal cancer risk in Morocco: evidence from a nationally representative case–control study. *Eur J Nutr*. 2022;61:2507–2515.
9 Romaguera D, Fernández-Barrés S, Gracia-Lavedán E, et al. Consumption of ultra-processed foods and drinks and colorectal, breast, and prostate cancer. *Clin Nutr*. 2021;40:1537–1545.
10 Jafari F, Yarmand S, Nouri M, et al. Ultra-processed food intake and risk of colorectal cancer: a matched case-control study. *Nutr Cancer*. 2023;75:532–541.
11 Al Nahas A, Yammine Ghantous S, Morales Berstein F, et al. Associations between degree of food processing and colorectal cancer risk in a large-scale European cohort. *Int J Cancer*. 2025. https://doi.org/10.1002/ijc.35361.
12 Wang F, Ugai T, Haruki K, et al. Healthy and unhealthy plant-based diets in relation to the incidence of colorectal cancer overall and by molecular subtypes. *Clin Transl Med*. 2022;12:e893–n/a.
13 B W, Rl Z, Qj O, Ym C, Yj F, Cx Z. Association of plant-based dietary patterns with the risk of colorectal cancer: a large-scale case-control study. *Food Funct*. 2022;13. https://doi.org/10.1039/d2fo01745h.
14 Kim J, Boushey CJ, Wilkens LR, Haiman CA, Le Marchand L, Park S-Y. Plant-based dietary patterns defined by a priori indices and colorectal cancer risk by sex and race/ethnicity: the multiethnic cohort study. *BMC Med*. 2022;20:430.
15 Hanley-Cook GT, Huybrechts I, Biessy C, et al. Food biodiversity and total and cause-specific mortality in 9 European countries: an analysis of a prospective cohort study. *PLoS Med*. 2021;18:e1003834.
16 Huybrechts I, Chimera B, Hanley-Cook GT, et al. Food biodiversity and gastrointestinal cancer risk in nine European countries: analysis within a prospective cohort study. *Eur J Cancer*. 2024;210:114258.
17 Fardet A, Rock E. How to protect both health and food system sustainability? A holistic 'global health'-based approach via the 3V rule proposal. *Public Health Nutr*. 2020;23:3028–3044.
18 Riboli E, Hunt K, Slimani N, et al. European prospective investigation into cancer and nutrition (EPIC): study populations and data collection. *Public Health Nutr*. 2002;5:1113–1124.

19 Huybrechts I, Rauber F, Nicolas G, et al. *Characterization of the degree of food processing in the european prospective investigation into cancer and nutrition: application of the nova classification and validation using selected biomarkers of food processing*. 2022. https://doi.org/10.3389/fnut.2023.1207555.
20 Monteiro CA, Cannon G, Levy RB, et al. Ultra-processed foods: what they are and how to identify them. *Public Health Nutr*. 2019;22:936–941.
21 Satija A, Bhupathiraju SN, Spiegelman D, et al. Healthful and unhealthful plant-based diets and the risk of coronary heart disease in U.S. adults. *J Am Coll Cardiol*. 2017;70:411–422.
22 Daly AJ, Baetens JM, De Baets B. Ecological diversity: measuring the unmeasurable. *Mathematics*. 2018;6:119.
23 Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol*. 2019;16:713–732.
24 Morales-Berstein F, Biessy C, Viallon V, et al. Ultra-processed foods, adiposity and risk of head and neck cancer and oesophageal adenocarcinoma in the european prospective investigation into cancer and nutrition study: a mediation analysis. *Eur J Nutr*. 2024;63:377–396.
25 Etemadi A, Abnet CC, Graubard BI, et al. Anatomical subsite can modify the association between meat and meat compounds and risk of colorectal adenocarcinoma: findings from three large US cohorts. *Int J Cancer*. 2018;143:2261–2270.
26 Murphy N, Moreno V, Hughes DJ, et al. Lifestyle and dietary environmental factors in colorectal cancer susceptibility. *Mol Aspects Med*. 2019;69:2–9.
27 Tan BL, Liew W-P-P, Norhaizan ME. Nutrients and oxidative stress: friend or foe? *Oxid Med Cell Longev*. 2018:1–24.
28 Kliemann N, Al Nahas A, Vamos EP, et al. Ultra-processed foods and cancer risk: from global food systems to individual exposures and mechanisms. *Br J Cancer*. 2022. https://doi.org/10.1038/s41416-022-01749-y.
29 Rauber F, da Costa Louzada ML, Steele EM, Millett C, Monteiro CA, Levy RB. Ultra-processed food consumption and chronic non-communicable diseases-related dietary nutrient profile in the UK (2008–2014). *Nutrients*. 2018;10. https://doi.org/10.3390/nu10050587.
30 Gehring J, Touvier M, Baudry J, et al. Consumption of ultra-processed foods by pesco-vegetarians, vegetarians, and vegans: associations with duration and age at diet initiation. *J Nutr*. 2021;151:120–131.
31 Sadler CR, Grassby T, Hart K, Raats M, Sokolović M, Timotijevic L. Processed food classification: conceptualisation and challenges. *Trends Food Sci Technol*. 2021;112:149–162.
32 Chen Z, Khandpur N, Desjardins C, et al. Ultra-processed food consumption and risk of type 2 diabetes: three large prospective U.S. cohort studies. *Diabetes Care*. 2023;46:1335–1344.
33 Hanley-Cook GT, Gie SM, Parraguez JP, et al. Cross-context equivalence and agreement of healthy diet metrics for national and global monitoring: a multicountry analysis of cross-sectional quantitative 24-hour dietary intake studies. *Am J Clin Nutr*. 2024;120:1093–1104.