Metabolic Signatures of Healthy Lifestyle Patterns and Colorectal Cancer Risk in a European Cohort

Joseph A. Rothwell,[*,‡,§] Neil Murphy,[§] Jelena Bešević,[||] Nathalie Kliemann,[§] Mazda Jenab,[§] Pietro Ferrari,[§] David Achaintre,[§] Audrey Gicquiau,[§] Béatrice Vozar,[§] Augustin Scalbert,[§] Inge Huybrechts,[§] Heinz Freisling,[§] Cornelia Prehn,[||] Jerzy Adamski,[¶,#,**] Amanda J. Cross,[||] Valeria Maria Pala,[||] Marie-Christine Boutron-Ruault,[*,‡] Christina C. Dahm,[‡‡] Kim Overvad,[‡‡] Inger Torhild Gram,[§§] Torkjel M. Sandanger,[§§] Guri Skeie,[§§] Paula Jakszyn,[||||,¶¶] Kostas K. Tsilidis,[||,##] Krasimira Aleksandrova,[***,‡‡‡] Matthias B. Schulze,[‡‡‡,§§§] David J. Hughes,[|||||] Bethany van Guelpen,[¶¶¶] Stina Bodén,[¶¶¶] Maria-José Sánchez,[|||||||,‡‡‡‡‡‡,§§§§§§] Julie A. Schmidt,[****] Verena Katzke,[‡‡‡‡] Tilman Kühn,[‡‡‡‡] Sandra Colorado-Yohar,[§§§§,|||||||,¶¶¶¶] Rosario Tumino,[####] Bas Bueno-de-Mesquita,[*****] Paolo Vineis,[||,‡‡‡‡‡] Giovanna Masala,[§§§§§] Salvatore Panico,[||||||||] Anne Kirstine Eriksen,[¶¶¶¶¶] Anne Tjønneland,[¶¶¶¶¶] Dagfinn Aune,[||,#####,******] Elisabete Weiderpass,[§] Gianluca Severi,[*,‡] Véronique Chajès,[§] and Marc J. Gunter[§]

*Centre for Epidemiology and Population Health, U1018, Generations and Health Team, Faculté de Médecine, Université Paris-Saclay, INSERM, Villejuif, France; ‡Gustave Roussy, Villejuif, France; §International Agency for Research on Cancer, Lyon, France; ||Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom; ¶Research Unit, Molecular Endocrinology and Metabolism, Helmholtz Zentrum München, Neuherberg, Germany; #Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; **Experimental Genetics, School of Life Science Weihenstephan, Technische Universität München, Neuherberg, Germany; ‡‡Department of Public Health, Aarhus University, Aarhus C, Denmark; §§Faculty of Health Sciences, Department of Community Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway; ||||Unit of Nutrition and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of Oncology, Barcelona, Spain; ¶¶Blanquerna School of Health Sciences, Ramon Llull University, Barcelona, Spain; ##Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece; ***Nutrition, Immunity and Metabolism Group, Department of Nutrition and Gerontology, §§§Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; ‡‡‡Institute of Nutritional Science, University of Potsdam, Potsdam, Germany; |||||School of Biomolecular and Biomedical Science, Conway Institute, University College Dublin, Dublin, Ireland; ¶¶¶Department of Radiation Sciences, Oncology Unit, Umea University, Umea, Sweden; ****Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom; ‡‡‡‡Division of Cancer Epidemiology, Deutsches Krebsforschungszentrum, Stiftung des Öffentlichen Rechts, Heidelberg, Germany; §§§§Department of Epidemiology, Murcia Regional Health Council, Instituto Murciano de Investigatión Biomédica (IMIB)–Arrixaca, Murcia, Spain; ||||||CIBER Epidemiología y Salud Pública, Spain; ¶¶¶¶Research Group on Demography and Health, National Faculty of Public Health, University of Antioquia, Medellín, Colombia; ####Cancer Registry and Histopathology Department, Provincial Health Authority, Ragusa, Italy; *****Department for Determinants of Chronic Diseases, National Institute for Public Health and the Environment, BA Bilthoven, The Netherlands; ‡‡‡‡‡Italian Institute of Technology, Genova, Italy; §§§§§Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network–Istituto per lo Studio, la Prevenzione e la Rete Oncologica (ISPRO), Florence, Italy; ||||||||Dipartimento di Medicina Clinica e Chirurgia, Federico II University, Naples, Italy; ¶¶¶¶¶Danish Cancer Society Research Center, Diet, Genes and Environment, Copenhagen, Denmark; #####Department of Nutrition, Bjørknes University College, Oslo, Norway; ******Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway; ‡‡‡‡‡‡Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain; §§§§§§Department of Preventive Medicine and Public Health, University of Granada, Granada, Spain

BACKGROUND & AIMS: **Colorectal cancer risk can be lowered by adherence to the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) guidelines. We derived metabolic signatures of adherence to these guidelines and tested their associations with colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition cohort.**

Abbreviations used in this paper: **BMI, body mass index; EPIC, European Prospective Investigation into Cancer and Nutrition; OCFA, odd chain fatty acid; OR, odds ratio; PC, phosphatidylcholine; PLSR, partial least-squares regression; SFA, saturated fatty acid; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.**

1542-3565
https://doi.org/10.1016/j.cgh.2020.11.045

Most current article

METHODS: **Scores reflecting adherence to the WCRF/AICR recommendations (scale, 1–5) were calculated from participant data on weight maintenance, physical activity, diet, and alcohol among a discovery set of 5738 cancer-free European Prospective Investigation into Cancer and Nutrition participants with metabolomics data. Partial least-squares regression was used to derive fatty acid and endogenous metabolite signatures of the WCRF/AICR score in this group. In an independent set of 1608 colorectal cancer cases and matched controls, odds ratios (ORs) and 95% CIs were calculated for colorectal cancer risk per unit increase in WCRF/AICR score and per the corresponding change in metabolic signatures using multivariable conditional logistic regression.**

RESULTS: **Higher WCRF/AICR scores were characterized by metabolic signatures of increased odd-chain fatty acids, serine, glycine, and specific phosphatidylcholines. Signatures were inversely associated more strongly with colorectal cancer risk (fatty acids: OR, 0.51 per unit increase; 95% CI, 0.29–0.90; endogenous metabolites: OR, 0.62 per unit change; 95% CI, 0.50–0.78) than the WCRF/AICR score (OR, 0.93 per unit change; 95% CI, 0.86–1.00) overall. Signature associations were stronger in male compared with female participants.**

CONCLUSIONS: **Metabolite profiles reflecting adherence to WCRF/AICR guidelines and additional lifestyle or biological risk factors were associated with colorectal cancer. Measuring a specific panel of metabolites representative of a healthy or unhealthy lifestyle may identify strata of the population at higher risk of colorectal cancer.**

Keywords: Colorectal Neoplasm; Risk Factors; World Cancer Research Fund/American Institute for Cancer Research Recommendations; Targeted Metabolomics.

Colorectal cancer is one of the most common neoplasms, with approximately 1.8 million new cases and 860,000 deaths reported worldwide in 2018.[1] Established risk factors for colorectal cancer include adiposity, smoking, adult attained height, and high intake of alcohol and red and processed meat, whereas physical activity and high intakes of whole grains, fish, and dairy products may protect against the disease.[2] Therefore, individuals may be able to minimize their risk of colorectal cancer by following a healthy lifestyle and many thousands of cases per year could be avoided.

The World Cancer Research Fund and American Institute for Cancer Research (WCRF/AICR) issues continuously updated recommendations on diet, physical activity, and weight management for the prevention of cancer, based on all available evidence.[3] At their core are healthy behaviors in relation to weight maintenance, physical activity, and intakes of red and processed meat, fruit and vegetables, fiber, and alcohol. A summary score has been developed to measure individual adherence to recommendations.[4] Higher scores have since been found to be associated with colorectal cancer risk[4–8] and cancer-specific and overall mortality.[6]

Unhealthy lifestyle behaviors and low WCRF/AICR scores may increase the risk of colorectal cancer through adverse effects upon systemic metabolism. Although tumorigenesis is promoted by adiposity, hyperinsulinemia, and chronic inflammation,[9] the systemic metabolic changes that precede or precipitate these physiological states remain unclear. To identify specific metabolite patterns associated with lifestyle factors and then to investigate whether they may play a role in colorectal cancer development, we used an extensive set of participants for whom targeted metabolomics and fatty acid data had been acquired within the European Prospective Investigation into Cancer and Nutrition cohort (EPIC). The objective of this analysis was first to characterize metabolic signatures of the WCRF/AICR score in a large group of cancer-free controls and to identify which compounds contributed to these signatures, and, second, to determine whether these metabolic signatures in prediagnostic blood samples were associated with subsequent colorectal cancer development.

Materials and Methods

The European Prospective Investigation Into Cancer Cohort and Collection of Data and Samples

EPIC is a multicenter prospective cohort that was established to investigate risk factors for cancer and other chronic diseases. More than 520,000 healthy subjects were enrolled between 1992 and 2000 from 23 EPIC administrative centers in 10 European countries. The collection of participant data and biospecimens has been described previously.[10] WCRF/AICR scores were calculated for all participants from recommendations on weight maintenance, physical activity, intake of food and drinks that promote weight gain, intake of plant-based foods, intake of animal-based foods, alcohol intake, and breastfeeding (Supplementary Table 1). Although the recommendations were updated in 2018,[11] we retained the scores previously calculated in EPIC.[4] These ranged from 0 to 6 for men and from 0 to 7 for women and were

grouped into quintiles for statistical modeling. The data and samples used were from all EPIC countries except Greece. Approval for the study was obtained from the International Agency for Research on Cancer and the ethical review boards of the participating institutes. All participants provided written informed consent.

Metabolomics Study Design

This analysis used a discovery set of 5738 cancer-free control participants, originating from several non-colorectal case–control studies nested within the EPIC cohort, to derive metabolic signatures of the WCRF/AICR score (ie, the linear combination of metabolites optimally related to the score). Fasted plasma and serum samples from the discovery set of controls were analyzed for either 34 fatty acids extracted from phospholipid fractions (n = 4239) or 155 endogenous metabolites assayed by the Biocrates Absolute*IDQ* P150/P180 Kit (n = 1741; Biocrates Life Sciences AG, Innsbruck, Austria). These 2 analyses are referred to as *fatty acids* and *endogenous metabolites* throughout this article. Metabolic signatures were determined separately for the 2 analyses by multivariate partial least-square regression (PLSR) models. Metabolite-predicted scores then were determined for each participant in the nested colorectal case–control study (n = 1608 cases and 1608 matched controls) for whom fatty acid or endogenous data were available, and these were regarded as the magnitude of the metabolic signature. All case–control participants had been analyzed for endogenous metabolites, while a subset of 438 cases and 438 matched controls additionally were analyzed for fatty acids. Associations between colorectal cancer risk and fatty acid signature, endogenous metabolic signature, and WCRF/AICR score then were tested separately in multivariable-adjusted models. The study design is illustrated in Figure 1.

Follow-up Evaluation for Colorectal Cancer Incidence

Incident cases of colorectal cancer were identified from health insurance records, contact with cancer and pathology registries, and the active follow-up evaluation of participants. Cases were defined using the International Classification of Diseases, 10th revision, and the International Classification of Diseases for Oncology, 2nd revision. Cases were incidence-density matched to cancer-free controls by age and year of sampling, sex, study center, follow-up time since blood collection, fasting status, and, when relevant, menopausal status and phase of menstrual cycle at blood collection.

Acquisition of Metabolomics Data

Saturated fatty acids (SFAs), monounsaturated fatty acids, polyunsaturated fatty acids, industrial trans fatty acids, and natural trans fatty acids were extracted from plasma phospholipid fractions and quantified by gas chromatography.[12] For endogenous metabolites, the Biocrates Absolute*IDQ* p150 or p180 Kits were used to measure concentrations of amino acids, biogenic amines, hexose sugars, acylcarnitines, sphingolipids (sphingomyelins), phosphatidylcholines (PC), and lysophosphatidylcholines in serum or plasma, following the recommended procedure.[13,14] See the Supplementary Methods section for further details of analytical methodology.

What You Need to Know

Background

The World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) score is a composite of diet and lifestyle variables and has been found to be associated inversely with colorectal cancer risk in previous studies.

Findings

Blood fatty acid and endogenous metabolite signatures of the WCRF/AICR score derived from a discovery set 5738 of cancer-free participants were associated more strongly with colorectal cancer risk than the WCRF/AICR score as calculated from baseline participant data in a study of 1608 colorectal cancer cases and 1608 matched controls.

Implications for patient care

Metabolic signatures of the WCRF/AICR score may capture etiologic risk factors for colorectal cancer beyond the score itself and provide insight into metabolic changes that precede cancer development. If replicated, measurement of these metabolite signatures could help identify strata of the population at higher risk of colorectal cancer.

Statistical Analysis

Determination of metabolic signatures. Discovery set metabolite data were $\log_2$ transformed, scaled, and missing values were imputed with minimum values. The resulting matrices were transformed to the residuals of a linear model on sex, batch, center (fixed effects), and study (random effects). Metabolic signatures were derived as the loadings (coefficients) on the first latent variable of a PLSR model (p_{LV1}) with metabolites as predictors and WCRF/AICR score as the response. The validated PLSR models then were used to predict WCRF/AICR scores in the case–control study on a continuous scale of 1 to 5. Pearson correlations between metabolite concentrations also were calculated in a subset of participants. See the Supplementary Methods section for further details.

Association of metabolic signatures of World Cancer Research Fund/American Institute for Cancer Research score with adherence to recommendations and colorectal

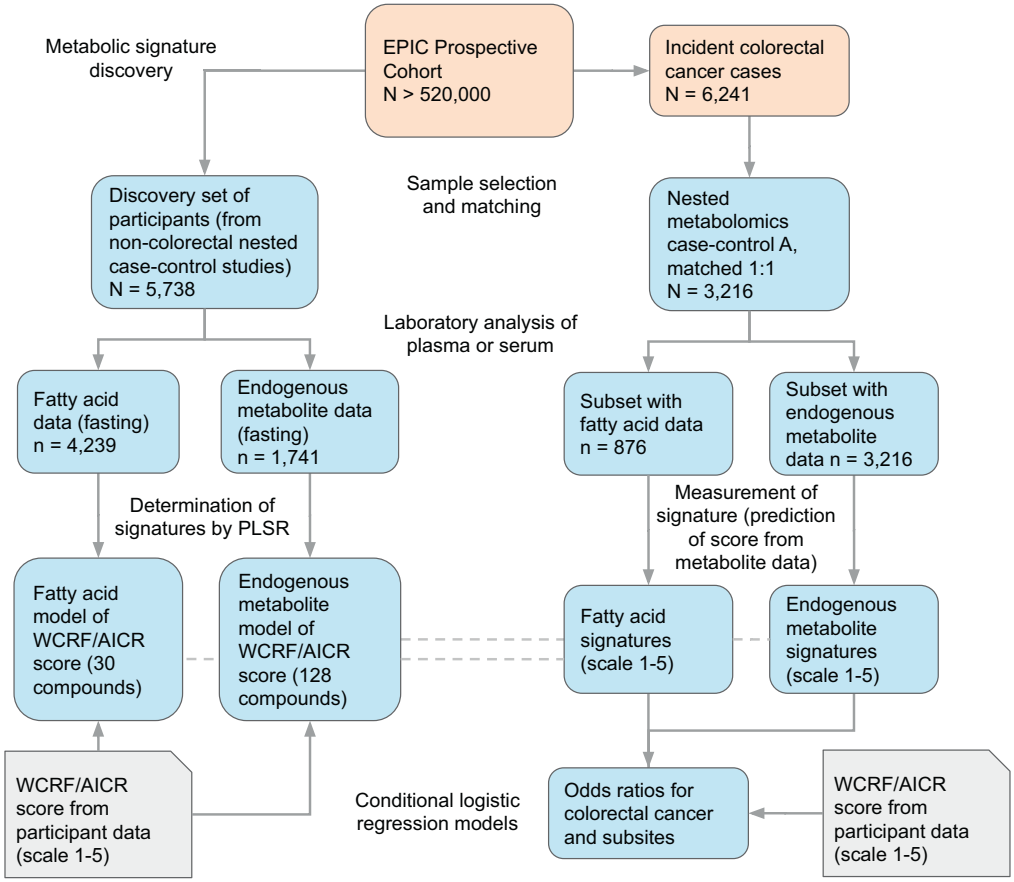


Figure 1. Overview of the study design. An independent set of healthy controls (*left*) was used to derive metabolic signatures of the WCRF/AICR score, which then were used to predict score categories in the nested case–control study (*right*). EPIC, European Prospective Investigation into Cancer and Nutrition; PLSR, partial least-squares regression; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.

cancer risk. Partial Pearson correlations were calculated between metabolic signatures and adherence to the 6 individual components of the WCRF/AICR score (as given earlier, each on a scale of 0, 0.5, or 1), adjusting for height, highest education level attained, and smoking status and intensity. Odds ratios and 95% CIs were calculated for risk of colorectal cancer and subsites with a metabolic signature or WCRF/AICR score as the main explanatory variable in multivariable conditional logistic regression models. Additional models were fit for individual WCRF/AICR components. Sensitivity analyses also were performed, additionally adjusting for smoking duration, intake of dairy products, or, in signature models only, WCRF/AICR score. Extra analyses were performed by strata of follow-up time and, for signatures only, body mass index (BMI) and WCRF/AICR score. All analyses were performed using R statistical software (Vienna, Austria), version 3.6.2.

Results

Characteristics of Nested Case–Control Study Participants

Participant characteristics for the nested case–control study are shown in Table 1. Cases were followed up for an average of 7.7 years before a colorectal cancer diagnosis. Cases had a higher BMI and larger waist circumference than controls at baseline, were taller, and attained lower WCRF/AICR scores. Participant characteristics for the discovery set are shown in Supplementary Table 2.

Metabolomics Data and Metabolic Signatures of World Cancer Research Fund/American Institute for Cancer Research Score

A total of 155 endogenous metabolites and 34 fatty acids were measured in both discovery and case–control data sets (Supplementary Table 3). Many high correlations ($r > 0.9$) were noted within metabolite classes (Supplementary Figure 1), but fewer were noted between compounds from fatty acid and endogenous metabolite platforms, with r greater than 0.6 for only 25 of 4964 possible correlations (Figure 2*A* and Supplementary Table 4). In the discovery set, the case–control study of origin contributed most variability to endogenous metabolite profiles with a partial R-square statistic ($R_{partial}^2$) of 20.3% (Supplementary Figure 2), while the study center explained most variability in fatty acid profiles ($R_{partial}^2 = 3.0\%$).

After exclusion of compounds with insufficient detection rates or high coefficient of variations, 128 endogenous compounds and 30 fatty acids remained for the derivation of metabolic signatures. Of these, SFAs 17:0 and SFAs 15:0 ($p_{LV1} = 0.149$ and 0.076, respectively) were increased most markedly in the fatty acid

Table 1. Characteristics of the Colorectal Cancer Cases and Matched Controls in EPIC

	Controls	Cases	*P* value[a]
N	1608	1608	
Sex			
Male	730 (45.4)	730 (45.4)	–
Female	878 (54.6)	878 (54.6)	
Age at blood collection, *y*	56.8 ± 7.5	56.9 ± 7.5	.74
Time to diagnosis, *y*	–	7.7 ± 4.4	–
Country			
France	52 (3.2)	52 (3.2)	–
Italy	387 (24.1)	387 (24.1)	
Spain	317 (19.7)	317 (19.7)	
United Kingdom	243 (15.1)	243 (15.1)	
The Netherlands	139 (8.6)	139 (8.6)	
Germany	163 (10.1)	163 (10.1)	
Denmark	307 (19.1)	307 (19.1)	
Tumor site			
Proximal colon	–	599 (37.7)	–
Distal colon	–	657 (41.3)	
Rectum	–	233 (14.7)	
Other	–	100 (6.3)	
Unknown	–	19 (1.2)	
Confirmed histologic verification			
Yes	–	1387 (86.3)	–
No	–	221 (13.7)	
Smoking status			.06
Nonsmoker	759 (47.2)	683 (42.5)	
Never smoker	480 (29.9)	519 (32.3)	
Smoker	353 (22.0)	390 (24.3)	
Height, *cm*	165.6 ± 9.3	166.1 ± 9.3	.008
BMI, kg/m^2	26.4 ± 3.9	27.0 ± 4.4	<.001
Waist circumference, *cm*	88.0 ± 12.2	90.4 ± 13.2	<.001
Total energy intake, *kcal*	2177 ± 643	2160 ± 702	.41
Physical activity, *MET*	87.7 ± 52.7	84.3 ± 52.6	.66
Alcohol intake, *g/d*	15.0 ± 18.9	16.7 ± 21.5	.09
WCRF/AICR score	2.54 ± 1.02	2.46 ± 1.02	.03
Fatty acid metabolic signature	2.64 ± 0.41	2.59 ± 0.42	<.001
Endogenous metabolic signature	2.51 ± 0.27	2.47 ± 0.30	.015

NOTE. Means and SD or frequency and percentage are shown unless stated otherwise.
BMI, body mass index; EPIC, European Prospective Investigation into Cancer and Nutrition cohort; MET, metabolic equivalent of task; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.
[a]*P* value for paired *t* test, Wilcoxon signed-rank test, or chi-squared test. Matching factors were age, sex, study center, follow-up time since blood collection, fasting status, menopausal status, and phase of menstrual cycle at blood collection.

signature of high WCRF/AICR scores (Table 2 and Figure 2*B*), while monounsaturated fatty acids 16:1n-7/n-9 and SFAs 16:0 were most diminished (p_{LV1} = -0.058 and -0.043, respectively). The endogenous metabolic signature of the WCRF/AICR score was dominated by phosphatidylcholines (PCs). Lysophosphatidylcholines a 17:0, PC ae 40:6 and PC ae C36:2 were most increased for high scores (p_{LV1} = 0.035, 0.032, and 0.032, respectively), while PC aa C32:1 and PC aa C38:4 were most diminished (p_{LV1} = -0.037 and -0.034, respectively).

Association Between Metabolic Signatures, World Cancer Research Fund/American Institute for Cancer Research Score Components and Colorectal Cancer Risk

Both metabolic signatures were correlated significantly with adherence to the weight maintenance and alcohol avoidance recommendations (Figure 2*C*). Fatty acid signatures captured the alcohol guideline to the greatest extent (r = 0.43) and endogenous metabolite weight maintenance (r = 0.33). A 1-unit increase in the fatty acid signature was associated with a 49% lower risk of colorectal cancer (odds ratio [OR], 0.51 per unit increase; 95% CI, 0.29–0.90), while a 1-unit increment in the endogenous metabolic signature (scale, 1–5) was associated with a 38% lower risk of colorectal cancer (OR, 0.62 per unit; 95% CI, 0.50–0.78). In comparison, a 1-unit increase in the WCRF/AICR score was associated with a 7% lower risk in the whole case–control study (OR, 0.93 per unit; 95% CI, 0.86–1.00) (Table 3). For comparison, associations between adherence to individual WCRF/AICR components and colorectal cancer risk are shown in Supplementary Table 5. By anatomic subsite, a 1-unit increment in the metabolic signature of endogenous metabolites was associated with a 35% lower risk of colon cancer (OR, 0.65 per unit; 95% CI, 0.50–0.84) and a 56% lower risk of rectal cancer (OR, 0.44 per unit; 95% CI, 0.25–0.79). As an additional analysis, when signature models additionally were adjusted for the WCRF/AICR score, the association between colorectal cancer risk and the fatty acid signature lost statistical significance (OR, 0.59 per unit; 95% CI, 0.33–1.07), whereas the association for the endogenous metabolic signature was not changed appreciably (OR, 0.62 per unit; 95% CI, 0.49–0.79). Sensitivity analyses are presented in Supplementary Table 6.

Discussion

In this analysis, we have derived fatty acid and endogenous metabolite signatures associated with the WCRF/AICR score from a large group of cancer-free control participants. Signatures were characterized by specific profiles of odd chain fatty acids (OCFAs), PCs, and amino acids, and principally captured the weight management and alcohol avoidance aspects of the WCRF/AICR guidelines. Both signatures were associated more strongly with colorectal cancer risk than the traditional WCRF/AICR score in the same participants. Measuring these signatures could provide a more sensitive assessment of colorectal cancer risk than

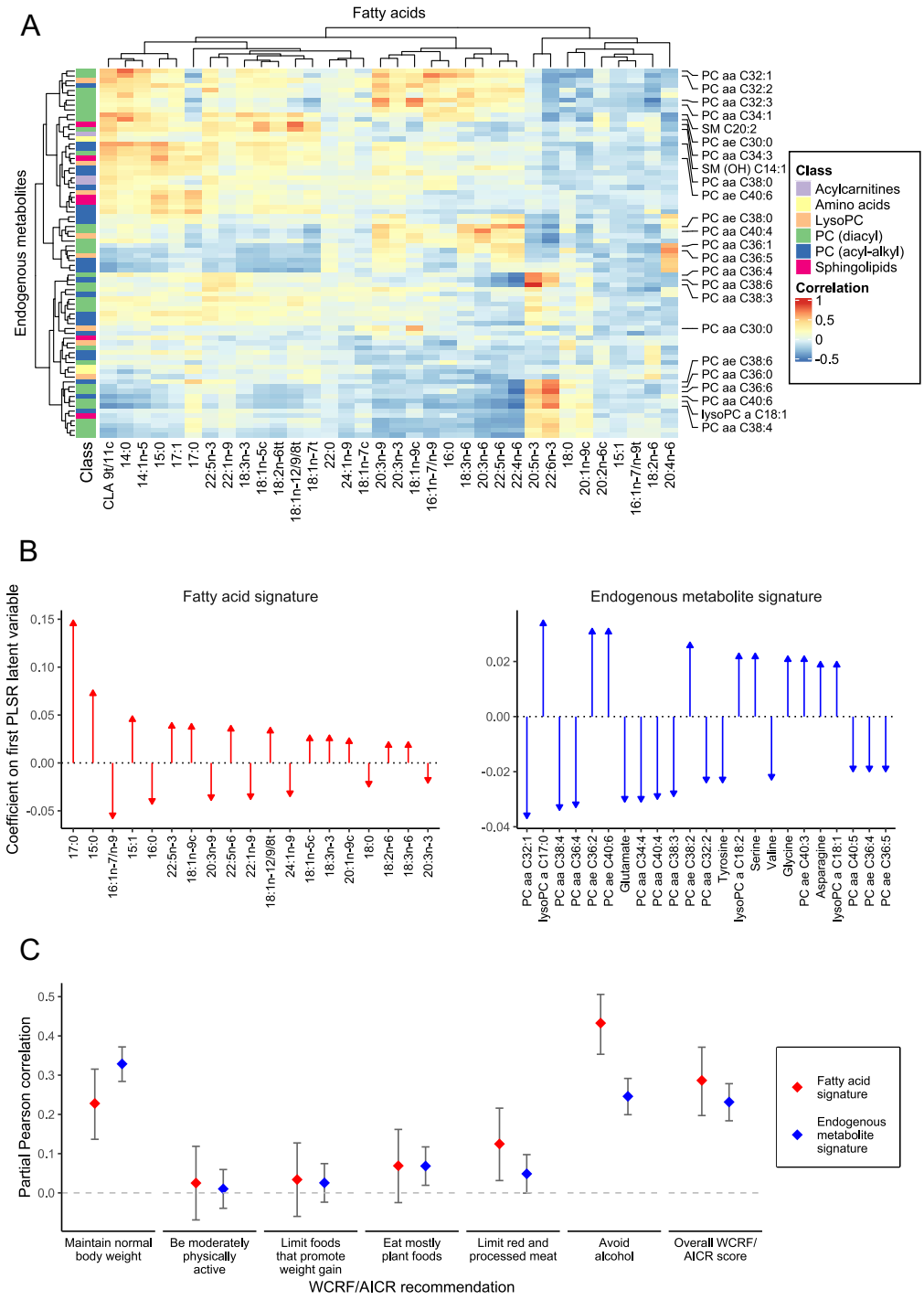


Figure 2. (*A*) Pearson correlations between fatty acids and endogenous metabolites in 439 control participants. Endogenous metabolites with no correlations greater than 0.25 with fatty acids have been omitted. (*B*) Strongest components of fatty acid and endogenous metabolite signatures of high WCRF/AICR scores in order of coefficient magnitude in PLSR models. (*C*) Partial correlations between individual WCRF/AICR recommendation scores and metabolic signatures in control participants. Partial correlations were adjusted for height, energy intake, highest educational level attained, smoking status, and smoking intensity. lyso PC, lysophosphatidylcholine; PC, phosphatidylcholine; PLSR, partial least-squares regression; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.

questionnaire data and physical measurements alone because they may encompass a greater range of lifestyle behaviors and characteristics than those captured by the WCRF/AICR recommendations.

Adherence to the WCRF/AICR guidelines has been associated with a reduced risk of colorectal cancer in EPIC and other cohorts. Previous studies have used custom weightings for score components; for example, to best capture colorectal cancer–specific risk factors.[7] We weighted score components evenly to characterize the metabolic profiles that accompany general cancer-preventing or cancer-promoting lifestyles. In terms of individual compounds, OCFA 17:0 and 15:0 were strikingly influential in the fatty acid signature. OCFAs originate from dairy fat and significant correlations between total OCFAs and dairy product intakes have been reported previously.[15,16] However, adjustment for total dairy product intake in our analysis changed risk estimates only minimally. Other factors also may affect circulating OCFAs, such as alcohol[16] and fiber intake via de novo formation from propionate.[17] OCFAs have also been positively associated with a lower incidence of type

Table 2. Compounds Contributing Most to Metabolic Signatures of WCRF/AICR Score by Coefficient in the First PLSR Latent Variable

Components of metabolic signature	Metabolite subclass or description	Coefficient from first LV of PLSR model, p_{LV1}[a]	OR (95% CI) for association with colorectal cancer[b]
Fatty acids[c]			
Increased for higher WCRF/AICR scores			
17:0	Saturated FA (odd chain)	0.149	**0.81 (0.71–0.99)**
15:0	Saturated FA (odd chain)	0.076	**0.78 (0.65–0.93)**
15:1	Monounsaturated FA	0.049	0.99 (0.85–1.16)
22:5n-6	Polyunsaturated FA	0.042	0.95 (0.80–1.13)
18:1n-9c	Monounsaturated FA	0.041	1.07 (0.92–1.26)
Diminished for higher WCRF/AICR scores			
16:1n-7/n-9	Monounsaturated FA	-0.058	0.96 (0.80–1.14)
16:0	Saturated FA	-0.043	0.92 (0.78–1.09)
20:3n-9	Polyunsaturated FA	-0.039	0.99 (0.84–1.17)
22:1n-9	Monounsaturated FA	-0.038	1.10 (0.91–1.32)
Endogenous metabolites[d]			
Increased for higher WCRF/AICR scores			
lysoPC a C17:0	Lysophosphatidylcholine	0.035	0.80 (0.62–1.02)
PC ae C40:6	Phosphatidylcholine, acyl-alkyl	0.032	0.90 (0.72–1.14)
PC ae C36:2	Phosphatidylcholine, acyl-alkyl	0.032	**0.72 (0.54–0.97)**
PC ae C38:2	Phosphatidylcholine, acyl-alkyl	0.027	0.90 (0.70–1.15)
Serine	Amino acid	0.023	0.87 (0.63–1.20)
lysoPC a C18:2	Lysophosphatidylcholine	0.023	0.85 (0.66–1.10)
Glycine	Amino acid	0.022	0.83 (0.62–1.13)
Diminished for higher WCRF/AICR scores			
PC aa C32:1	Phosphatidylcholine, diacyl	-0.037	0.94 (0.72–1.23)
PC aa C38:4	Phosphatidylcholine, diacyl	-0.034	1.13 (0.89–1.42)
PC aa C36:4	Phosphatidylcholine, diacyl	-0.033	1.08 (0.83–1.39)
Glutamate	Amino acid	-0.031	1.12 (0.64–1.97)
PC aa C34:4	Phosphatidylcholine, diacyl	-0.031	0.83 (0.66–1.06)
PC aa C40:4	Phosphatidylcholine, diacyl	-0.030	1.04 (0.83–1.30)
PC ae C38:3	Phosphatidylcholine, acyl-alkyl	-0.029	0.79 (0.61–1.02)

NOTE. Boldface indicates statistical significance.
CI, confidence interval; FA, fatty acid; LV, latent variable; lysoPC, lysophosphatidylcholine; OR, odds ratio; PC, phosphatidylcholine; PLSR, partial least-squares regression; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.
[a]After adjustment for center, batch, and study using the residuals method. Coefficients for all compounds are shown in Supplementary Table 3.
[b]Fatty acids, odds ratio per SD increase in concentration; endogenous metabolites, odds ratio for fourth vs first quartile of compound concentration. Adjusted for body mass index, alcohol intake, red and processed meat intake, height, energy intake, highest educational level attained, smoking status, and smoking intensity.
[c]Compounds with coefficients in the top or bottom quintiles for the first PLSR LV.
[d]Compounds with coefficients in the top or bottom 5 percentiles for the first PLSR LV.

2 diabetes[18] and an anti-inflammatory profile of adipokines.[19] Fatty acid intake is known to modulate biomarkers of inflammation.[20]

Fatty acids obtained from the diet also are incorporated into PCs, which are components of biological membranes but also signaling molecules that govern processes such as gene regulation and homeostatic control of serum glucose.[21] PCs that are influential in the endogenous metabolite signature have been linked to individual lifestyle behaviors in previous studies. LysoPC a C17:0 and PC ae C36:2, increased in the signature of a high WCRF/AICR score, were associated inversely with alcohol intake in 3 separate prospective studies.[22,23] PC aa C32:1, conversely, was associated positively with alcohol intake in the same studies, and associated independently with high total meat intake, smoking, and risk of type 2 diabetes.[24–27] Since PCs are readily perturbed by diet and lifestyle factors and fine differences in structure impart distinct bioactivities, dedicated studies are needed to elucidate their relationship to tumorigenesis. Glycine, increased in the endogenous signature of a high WCRF/AICR score, has been reported to be associated inversely with total red meat intake[28] and type 2 diabetes risk,[26] but associated positively with total weekly physical activity.[29] Glutamate, conversely, appeared in metabolic profiles of a high BMI[30] and is associated with insulin resistance.[31] Our observations regarding amino acids were largely consistent with previous studies.

Both signatures captured weight management and alcohol avoidance more strongly than other components of the WCRF/AICR score, despite the orthogonality of the 2 platforms. Alcohol avoidance was captured strikingly by the fatty acid signature. OCFAs in particular have been reported to be associated inversely with alcohol intake,[16,32] although ethanol exposure may attenuate fatty acid absorption and incorporation into phospholipids by diverse mechanisms such as inhibition of

Table 3. ORs and 95% CI for Colorectal Cancer Risk and Metabolic Signatures or WCRF/AICR Score by Sex and Anatomic Subsite

	Colorectal OR (95% CI)	Colon OR (95% CI)	Proximal colon OR (95% CI)	Distal colon OR (95% CI)	Rectal OR (95% CI)
	N = 3216	N = 2504	N = 1190	N = 1314	N = 468
Fatty acids					
N, women	876 (530)	792 (486)	358 (226)	434 (260)	
WCRF/AICR score[a]					
All	**0.77 (0.66–0.91)**	**0.75 (0.63–0.89)**	0.83 (0.63–1.10)	**0.70 (0.55–0.90)**	
Women	**0.78 (0.63–0.98)**	**0.77 (0.61–0.97)**	0.87 (0.58–1.29)	0.73 (0.53–1.01)	
Men	**0.75 (0.58–0.96)**	**0.69 (0.52–0.92)**	0.74 (0.48–1.15)	0.64 (0.42–0.97)	
P het	.36	.28	.44	.49	
Metabolic signature[a,b]					
All	**0.51 (0.29–0.90)**	**0.53 (0.29–0.97)**	0.78 (0.31–1.97)	**0.40 (0.18–0.91)**	
Women	0.73 (0.34–1.57)	0.77 (0.34–1.71)	0.67 (0.18–2.44)	0.70 (0.24–2.00)	
Men	**0.31 (0.13–0.75)**	**0.33 (0.13–0.83)**	0.84 (0.18–4.00)	**0.23 (0.06–0.83)**	
P het	.072	.11	.43	.18	
Metabolic signature adjusted for WCRF/AICR score					
All	0.59 (0.33–1.07)	0.61 (0.33–1.14)	0.79 (0.30–2.02)	0.52 (0.22–1.21)	
Endogenous metabolites					
N, women	3216 (1752)	2504 (1418)	1190 (712)	1314 (706)	468 (258)
WCRF/AICR score[a]					
All	**0.93 (0.86–1.00)**	0.93 (0.85–1.02)	1.00 (0.87–1.14)	0.89 (0.79–1.01)	0.89 (0.72–1.08)
Women	1.01 (0.91–1.12)	1.05 (0.93–1.18)	1.07 (0.90–1.29)	1.04 (0.87–1.23)	0.96 (0.70–1.31)
Men	**0.85 (0.76–0.95)**	**0.80 (0.70–0.92)**	0.90 (0.73–1.12)	**0.72 (0.59–0.87)**	0.83 (0.62–1.11)
P het	.022	.002	.12	.005	.83
Metabolic signature[a,b]					
All	**0.62 (0.50–0.78)**	**0.65 (0.50–0.84)**	0.78 (0.53–1.14)	**0.57 (0.40–0.82)**	**0.44 (0.25–0.79)**
Women	0.82 (0.59–1.12)	0.89 (0.62–1.26)	0.92 (0.55–1.54)	0.87 (0.52–1.43)	0.60 (0.25–1.46)
Men	**0.44 (0.32–0.61)**	**0.44 (0.25–0.79)**	0.59 (0.33–1.06)	**0.36 (0.21–0.62)**	**0.41 (0.19–0.86)**
P het	.029	.03	.21	.12	.46
Metabolic signature adjusted for WCRF/AICR score					
All	**0.62 (0.49–0.79)**	**0.63 (0.48–0.83)**	**0.61 (0.42–0.90)**	**0.67 (0.45–1.00)**	**0.52 (0.29–0.94)**

NOTE. Boldface indicates statistical significance.
CI, confidence interval; OR, odds ratio; P het, P heterogeneity; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.
[a]On a scale of 1 to 5, and after adjustment for height, energy intake, highest educational level attained, and smoking status or intensity.
[b]Magnitude of the metabolic signature is defined as the metabolite-predicted WCRF/AICR score derived from partial least-squares regression models trained on endogenous metabolite and fatty acid data from the discovery set.

enzyme catalysts, disruption of gut microbiota, or physiological changes to hepatocytes.[23,33] Weight management was captured most strongly by the endogenous signature, whose amino acid components are implicated in adiposity and insulin resistance. In sensitivity analysis, the endogenous signature remained associated strongly with colorectal cancer risk after additional adjustment for the WCRF/AICR score, showing a capability to capture intrinsic or longer-term abnormalities in metabolism related to the disease. The fact that associations for metabolic signatures were stronger than those of WCRF/AICR scores suggests that signatures, rather than acting as biomarker surrogates of score, reflect aspects of metabolic health that are not measured directly by conventional approaches.[34]

The association of the metabolic signatures with colorectal cancer was more apparent in men and the associations were weaker and nonsignificant in women. This may reflect sex-specific differences in the association of the composite risk factors within the score such as BMI and alcohol consumption, which are stronger risk factors for colorectal cancer in men than in women.[35] In addition to this heterogeneity, it is known that colorectal cancer risk factors and associations by sex may differ by anatomic subsite,[36] and in our study associations for colon cancer were driven disproportionately by distal tumors. Interestingly, rectal cancer, however, was associated strongly with endogenous metabolic signatures of the WCRF/AICR score, despite the influence of biologic, lifestyle, and dietary factors upon risk being less clear than for colon cancer.[37] Overall, these differences require follow-up evaluation in other cohorts, but if reproduced may point toward specific biological pathways that deserve mechanistic investigation.

Our study is unique in deriving metabolic signatures from a large fasting discovery group on 2 complementary

platforms and measuring their magnitude prospectively in a nested case–control study of substantial size. One limitation is that we have been unable to test these signatures in external cohorts to date. Participants nonetheless were from different combinations of EPIC centers and samples were analyzed in different laboratories. Because endogenous metabolite and fatty acid data were not always available for the same participants, an overall signature derived from both platforms could not be determined, and the fatty acid signature was derived from a data set of mostly female participants and therefore may have been less applicable to males. Another drawback was the unavailability of data on colorectal cancer screening and family history and use of nonsteroidal anti-inflammatory drugs in some EPIC centers, meaning we were unable to adjust for these potential confounders.

In conclusion, the stronger associations of signatures with colorectal cancer compared with the WCRF/AICR scores suggest that metabolite profiles reflect a broader spectrum of behavioral and biological characteristics than are included in the recommendations and can be used to better assess colorectal cancer risk or gain insight into metabolic risk factors. Further studies of healthy lifestyle patterns and their relationship with metabolism and cancer are merited.

Supplementary Material

Note: To access the supplementary material accompanying this article, please click here.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68:394–424.
2. Continuous Update Project Expert Report 2018. Diet, nutrition, physical activity and colorectal cancer. 2018. Available at: dietandcancerreport.org. Accessed July 7, 2020.
3. Clinton SK, Giovannucci EL, Hursting SD. The World Cancer Research Fund/American Institute for Cancer Research third expert report on diet, nutrition, physical activity, and cancer: impact and future directions. J Nutr 2019;150:663–671.
4. Romaguera D, Vergnaud AC, Peeters PH, et al. Is concordance with World Cancer Research Fund/American Institute for Cancer Research guidelines for cancer prevention related to subsequent risk of cancer? Results from the EPIC study. Am J Clin Nutr 2012;96:150–163.
5. Turati F, Bravi F, Di Maso M, et al. Adherence to the World Cancer Research Fund/American Institute for Cancer Research recommendations and colorectal cancer risk. Eur J Cancer 2017;85:86–94.
6. Solans M, Chan DSM, Mitrou P, et al. A systematic review and meta-analysis of the 2007 WCRF/AICR score in relation to cancer-related health outcomes. Ann Oncol 2020; 31:352–368.
7. Petimar J, Smith-Warner SA, Rosner B, et al. Adherence to the World Cancer Research Fund/American Institute for Cancer Research 2018 recommendations for cancer prevention and risk of colorectal cancer. Cancer Epidemiol Biomarkers Prev 2019;28:1469–1479.
8. Hastert TA, White E. Association between meeting the WCRF/AICR cancer prevention recommendations and colorectal cancer incidence: results from the VITAL cohort. Cancer Causes Control 2016;27:1347–1359.
9. Murphy N, Jenab M, Gunter MJ. Adiposity and gastrointestinal cancers: epidemiology, mechanisms and future directions. Nat Rev Gastroenterol Hepatol 2018;15:659–670.
10. Riboli E, Hunt KJ, Slimani N, et al. European prospective investigation into cancer and nutrition (EPIC): study populations and data collection. Public Health Nutrition 2002; 5:1113–1124.
11. Shams-White MM, Brockton NT, Mitrou P, et al. Operationalizing the 2018 World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) cancer prevention recommendations: a standardized scoring system. Nutrients 2019;11:1572.
12. Chajes V, Assi N, Biessy C, et al. A prospective evaluation of plasma phospholipid fatty acids and breast cancer risk in the EPIC study. Ann Oncol 2017;28:2836–2842.
13. Stepien M, Duarte-Salles T, Fedirko V, et al. Alteration of amino acid and biogenic amine metabolism in hepatobiliary cancers: findings from a prospective cohort study. Int J Cancer 2016; 138:348–360.
14. Romisch-Margl W, Prehn C, Bogumil R, et al. Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. Metabolomics 2012;8:133–142.
15. Hodson L, Skeaff CM, Fielding BA. Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. Prog Lipid Res 2008;47:348–380.
16. Forouhi NG, Koulman A, Sharp SJ, et al. Differences in the prospective association between individual plasma phospholipid saturated fatty acids and incident type 2 diabetes: the EPIC-InterAct case-cohort study. Lancet Diabetes Endocrinol 2014;2:810–818.
17. Weitkunat K, Schumann S, Nickel D, et al. Odd-chain fatty acids as a biomarker for dietary fiber intake: a novel pathway for endogenous production from propionate. Am J Clin Nutr 2017;105:1544–1551.
18. Imamura F, Sharp SJ, Koulman A, et al. A combination of plasma phospholipid fatty acids and its association with incidence of type 2 diabetes: the EPIC-InterAct case-cohort study. PLoS Med 2017;14:e1002409.
19. Kurotani K, Sato M, Yasuda K, et al. Even- and odd-chain saturated fatty acids in serum phospholipids are differentially associated with adipokines. PLoS One 2017;12:14.
20. Baer DJ, Judd JT, Clevidence BA, et al. Dietary fatty acids affect plasma markers of inflammation in healthy men fed controlled diets: a randomized crossover study. Am J Clin Nutr 2004;79:969–973.
21. Furse S, de Kroon A. Phosphatidylcholine's functions beyond that of a membrane brick. Mol Membr Biol 2015;32:117–119.
22. van Roekel EH, Trijsburg L, Assi N, et al. Circulating metabolites associated with alcohol intake in the European Prospective Investigation into Cancer and Nutrition Cohort. Nutrients 2018;10:654.
23. Jaremek M, Yu Z, Mangino M, et al. Alcohol-induced metabolomic differences in humans. Transl Psychiatry 2013;3:8.
24. Schmidt JA, Rinaldi S, Ferrari P, et al. Metabolic profiles of male meat eaters, fish eaters, vegetarians, and vegans from the EPIC-Oxford cohort. Am J Clin Nutr 2015;102:1518–1526.
25. Wang-Sattler R, Yu Y, Mittelstrass K, et al. Metabolic profiling reveals distinct variations linked to nicotine consumption in

humans - first results from the KORA study. PLoS One 2008; 3:10.
26. Floegel A, Stefan N, Yu ZH, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. Diabetes 2013;62:639–648.
27. Assi N, Gunter MJ, Thomas DC, et al. Metabolic signature of healthy lifestyle and its relation with risk of hepatocellular carcinoma in a large European cohort. Am J Clin Nutr 2018;108:117–126.
28. Wittenbecher C, Muhlenbruch K, Kroger J, et al. Amino acids, lipid metabolites, and ferritin as potential mediators linking red meat consumption to type 2 diabetes. Am J Clin Nutr 2015; 101:1241–1250.
29. Ding M, Zeleznik OA, Guasch-Ferre M, et al. Metabolome-wide association study of the relationship between habitual physical activity and plasma metabolite levels. Am J Epidemiol 2019;188:1932–1943.
30. Carayol M, Leitzmann MF, Ferrari P, et al. Blood metabolic signatures of body mass index: a targeted metabolomics study in the EPIC cohort. J Proteome Res 2017;16:3137–3146.
31. Greenfield JR, Farooqi IS, Keogh JM, et al. Oral glutamine increases circulating glucagon-like peptide 1, glucagon, and insulin concentrations in lean, obese, and type 2 diabetic subjects. Am J Clin Nutr 2009;89:106–113.
32. Rosell M, Johansson G, Berglund L, et al. The relation between alcohol intake and physical activity and the fatty acids 14:0, 15:0 and 17:0 in serum phospholipids and adipose tissue used as markers for dairy fat intake. Br J Nutr 2005;93:115–121.
33. Irwin C, Van Reenen M, Mason S, et al. The H-1-NMR-based metabolite profile of acute alcohol consumption: a metabolomics intervention study. PLoS One 2018;13:20.
34. Murphy N, Cross AJ, Abubakar M, et al. A nested case-control study of metabolically defined body size phenotypes and risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC). PLoS Med 2016;13:e1001988.
35. Aleksandrova K, Pischon T, Jenab M, et al. Combined impact of healthy lifestyle factors on colorectal cancer: a large European cohort study. BMC Med 2014;12:168.
36. Murphy N, Ward HA, Jenab M, et al. Heterogeneity of colorectal cancer risk factors by anatomical subsite in 10 European countries: a multinational cohort study. Clin Gastroenterol Hepatol 2019;17:1323–1331.e6.
37. Wei EK, Giovannucci E, Wu K, et al. Comparison of risk factors for colon and rectal cancer. Int J Cancer 2004;108:433–442.

Reprint requests
Address requests for reprints to: Joseph A. Rothwell, PhD, Centre for Epidemiology and Population Health, U1018, Generations and Health Team, Faculté de Médecine, Université Paris-Saclay, INSERM, Villejuif, France. e-mail: joseph.rothwell@inserm.fr; fax: (33) 01-42-11-61-85.

Acknowledgments
The authors thank the European Prospective Investigation into Cancer and Nutrition centers of Norfolk (United Kingdom), Guipuzcoa (Spain), and Navarro (Spain) for provision of biological samples and participant data. The research was made possible using the data/samples provided by the International Agency for Research on Cancer Biobank.

CRediT Authorship Contributions
Joseph Rothwell, PhD (Conceptualization: Supporting; Formal analysis: Lead; Writing – original draft: Lead; Writing – review & editing: Lead);
Neil Murphy, PhD (Writing – review & editing: Equal);
Jelena Bešević, PhD (Writing – review & editing: Equal);
Nathalie Kliemann, PhD (Writing – review & editing: Equal);
Mazda Jenab, PhD (Writing – review & editing: Equal);
Pietro Ferrari, PhD (Writing – review & editing: Equal);
David Achaintre, PhD (Formal analysis: Equal);
Audrey Gicquiau, MSc (Formal analysis: Equal);
Béatrice Vozar, BTS (Formal analysis: Equal);
Augustin Scalbert, PhD (Writing – review & editing: Equal);
Inge Huybrechts, PhD (Writing – review & editing: Equal);
Heinz Freisling, PhD (Writing – review & editing: Equal);
Cornelia Prehn (Formal analysis: Equal; Writing – review & editing: Equal);
Jerzy Adamski (Writing – review & editing: Equal);
Amanda Cross, PhD (Writing – review & editing: Equal);
Valeria Pala, PhD (Writing – review & editing: Equal);
Marie-Christine Boutron-Ruault, PhD (Writing – review & editing: Equal);
Christina Dahm, PhD (Writing – review & editing: Equal);
Kim Overvad, PhD (Writing – review & editing: Equal);
Inger Torhild Gram, PhD (Writing – review & editing: Equal);
Torkjel Sandanger, PhD (Writing – review & editing: Equal);
Guri Skeie, PhD (Writing – review & editing: Equal);
Paula Jakszyn, PhD (Writing – review & editing: Equal);
Kostas Tsilidis, PhD (Writing – review & editing: Equal);
Krasimira Aleksandrova, PhD (Writing – review & editing: Equal);
Matthias Schulze, PhD (Writing – review & editing: Equal);
David Hughes, PhD (Writing – review & editing: Equal);
Bethany van Guelpen, PhD (Writing – review & editing: Equal);
Stina Bodén, MSc (Writing – review & editing: Equal);
Maria-José Sánchez, PhD (Writing – review & editing: Equal);
Julie Schmidt, PhD (Writing – review & editing: Equal);
Verena Katzke, PhD (Writing – review & editing: Equal);
Tilman Kühn, PhD (Writing – review & editing: Equal);
Sandra Colorado-Yohar, PhD (Writing – review & editing: Equal);
Rosario Tumino, PhD (Writing – review & editing: Equal);
Bas Bueno-de-Mesquita, PhD (Writing – review & editing: Equal);
Paolo Vineis, PhD (Writing – review & editing: Equal);
Giovanna Masala, PhD (Writing – review & editing: Equal);
Salvatore Panico, PhD (Writing – review & editing: Equal);
Anne Kirstine Eriksen, MSc (Writing – review & editing: Equal);
Anne Tjønneland, PhD (Writing – review & editing: Equal);
Dagfinn Aune, PhD (Writing – review & editing: Equal);
Elisabete Weiderpass, PhD (Writing – review & editing: Equal);
Gianluca Severi, PhD (Writing – review & editing: Equal);
Véronique Chajès, PhD (Supervision: Equal; Writing – review & editing: Equal);
Marc Gunter, PhD (Conceptualization: Lead; Funding acquisition: Lead; Supervision: Equal)

Conflicts of interest
The authors disclose no conflicts.

Funding
This work was supported by the World Cancer Research Fund grant 2013-1002 and the European Commission grant EU-FP7/BBMRI-LPC - number 313010 (M.J.G.). The coordination of the European Prospective Investigation into Cancer and Nutrition is supported financially by the International Agency for Research on Cancer and Imperial College London. The national cohorts are supported by the Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave-Roussy, Mutuelle Générale de l'Education Nationale, and Institut National de la Santé et de la Recherche Médicale (France); German Cancer Aid, German Cancer Research Center, Federal Ministry of Education and Research, Deutsche Krebshilfe, and Deutsches Krebsforschungszentrum (Germany); Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro Italy and the National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports, Netherlands Cancer Registry, LK Research Funds, Dutch Prevention Funds, Dutch Zorg Onderzoek Nederland, World Cancer Research Fund, and Statistics Netherlands (The Netherlands); Nordic Centre of Excellence Programme on Food, Nutrition and Health (Norway); Health Research Fund, PI13/00061 to Granada; PI13/01162 to European Prospective Investigation into Cancer and Nutrition–Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia (6236), and Navarra, Instituto de Salud Carlos III Cooperative Research in Health (RD06/0020) (Spain); Swedish Cancer Society, Swedish Research Council, and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to European Prospective Investigation into Cancer and Nutrition–Norfolk, and C570/A16491 and C8221/A19170 to European Prospective Investigation into Cancer and Nutrition–Oxford), and the Medical Research Council (1000143 to European Prospective Investigation into Cancer and Nutrition–Norfolk, and MR/M012190/1 to European Prospective Investigation into Cancer and Nutrition–Oxford) (United Kingdom). For information on how to submit an application for gaining access to European Prospective Investigation into Cancer and Nutrition data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php.

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.

Supplementary Methods

Laboratory Methods

Serum and plasma samples were stored at the International Agency for Research on Cancer (Lyon, France) at -196°C in liquid nitrogen, apart from those of Sweden (-80°C in freezers) and Denmark (-150°C in nitrogen vapor). Data and the biospecimens used were from all EPIC countries except Greece.

Fatty acid profiling was performed at the International Agency for Research on Cancer for both discovery and case–control samples. SFAs, monounsaturated fatty acids, polyunsaturated fatty acids, industrial trans fatty acids, and natural trans fatty acids were extracted from plasma phospholipid fractions and quantified using an Agilent 7890 gas chromatograph instrument (Agilent Technologies, Santa Clara, CA). Concentrations were expressed as the percentage of total fatty acids. For endogenous metabolites, analyses were performed at the International Agency for Research on Cancer (all discovery and approximately one third of case–control samples), and the Helmholtz Zentrum, München, Germany (all other case–control samples). The Absolute*IDQ* p150 or p180 Kits were used to measure concentrations of amino acids, biogenic amines, hexose sugars, acylcarnitines, sphingolipids, PCs, and lysoPCs in serum or plasma, following the recommended procedure. The International Agency for Research on Cancer method used a 1290 Series liquid chromatography instrument with a Q-Trap 5500 mass spectrometer (Agilent Technologies, Les Ulis, France). The Helmholtz method was based on a 1200 series liquid chromatography instrument (Agilent, Böblingen, Germany) with an API 4000 triple–quadrupole mass spectrometer (AB Sciex, Darmstadt, Germany). Case–control pairs were analyzed in the same batch, and coefficients of variation were calculated for each metabolite. The full details of the laboratory procedures have been published.[1–3]

Statistical Analysis

Determination of metabolic signatures. This analysis used a discovery set of 5738 cancer-free control participants, originating from several noncolorectal case–control studies nested within the EPIC cohort,[1,4–6] to derive metabolic signatures of the WCRF/AICR score. Discovery set metabolite matrices were prepared for derivation of metabolic signatures, separately for fatty acids and endogenous metabolites. Compounds not measured in both discovery and case–control sets were excluded, as well as those that were missing (outside the limits of quantification) for more than 40% of participants. For the remainder, missing concentrations were replaced with half the minimum in the whole data set. The discovery metabolite matrices then were $\log_2$ transformed, centered, and unit variance–scaled. Second, unwanted variability was removed from the data. The principal component partial R-squared technique was used to identify covariates that contributed the most toward variability in metabolomics data. The principal component partial R-squared technique combines principal component analysis and multivariable regression to estimate the relative effects of metadata variables upon a matrix of omics measurements.[7] Each metabolite concentration then was transformed by the residuals method[8] using models on sex, batch, center (fixed effects), and study (random effects). Pearson correlations between concentrations also were calculated in a subset of participants.

PLSR was used to determine metabolic signatures of the WCRF/AICR score[6] (ie, the linear combination of metabolite concentrations most correlated with the score). Models were selected that balanced simplicity and low root mean square error of cross-validation. Loadings (coefficients) on the first latent variable of the PLSR model fit, denoted p_{LV1}, were calculated for each compound as a measure of contribution to each signature. ORs and 95% CIs were calculated for colorectal cancer risk for baseline concentrations of compounds that contributed the most to these signatures, adjusting for BMI, height, energy intake, highest educational level attained, red and processed meat intake, alcohol intake, smoking status, and smoking intensity in conditional logistic regression models.

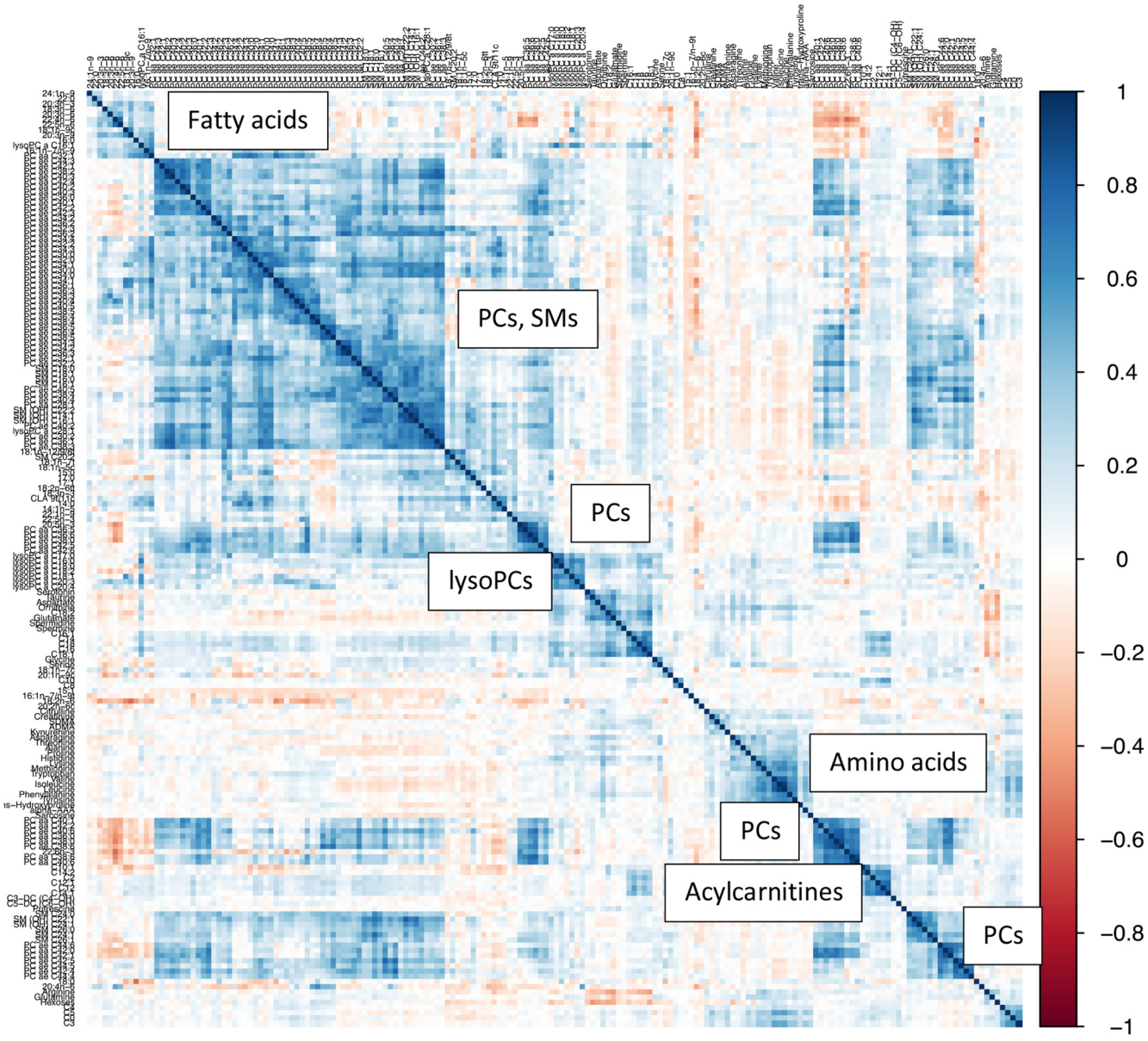
The case–control metabolite matrix was prepared similarly to that of the discovery set. The validated PLSR models then were used to predict WCRF/AICR scores, applying coefficients to metabolites, on a continuous scale of 1 to 5 for each subject in the case–control study. These predicted scores were regarded as the magnitude of the metabolic signature with distributions comparable with those of WCRF/AICR scores.

Association of metabolic signatures of the World Cancer Research Fund/American Institute for Cancer Research score with adherence to recommendations and colorectal cancer risk. Partial Pearson correlations were calculated between metabolic signatures and adherence to the 6 individual components of the WCRF/AICR score (as described earlier, each on a scale of 0, 0.5, or 1), adjusting for height, highest education level attained, smoking status, and intensity. Odds ratios and 95% CIs were calculated for risk of colorectal cancer and subsites, with the metabolic signature or the WCRF/AICR score as the main explanatory variable in multivariable conditional logistic regression models. Heterogeneity by sex was determined by likelihood ratio test, comparing unpaired logistic regression models with and without interaction terms between sex and the WCRF/AICR score or metabolic signature. Matching factors additionally were included in these models. Additional models were fit for individual WCRF/AICR components. Sensitivity analyses also were performed, additionally adjusting for smoking duration, intake of dairy products, or, in signature models only, WCRF/AICR score. Subgroup analyses

were performed for strata of follow-up time and, for signature only, BMI and WCRF/AICR score. All analyses were performed using R statistical software, version 3.6.2.

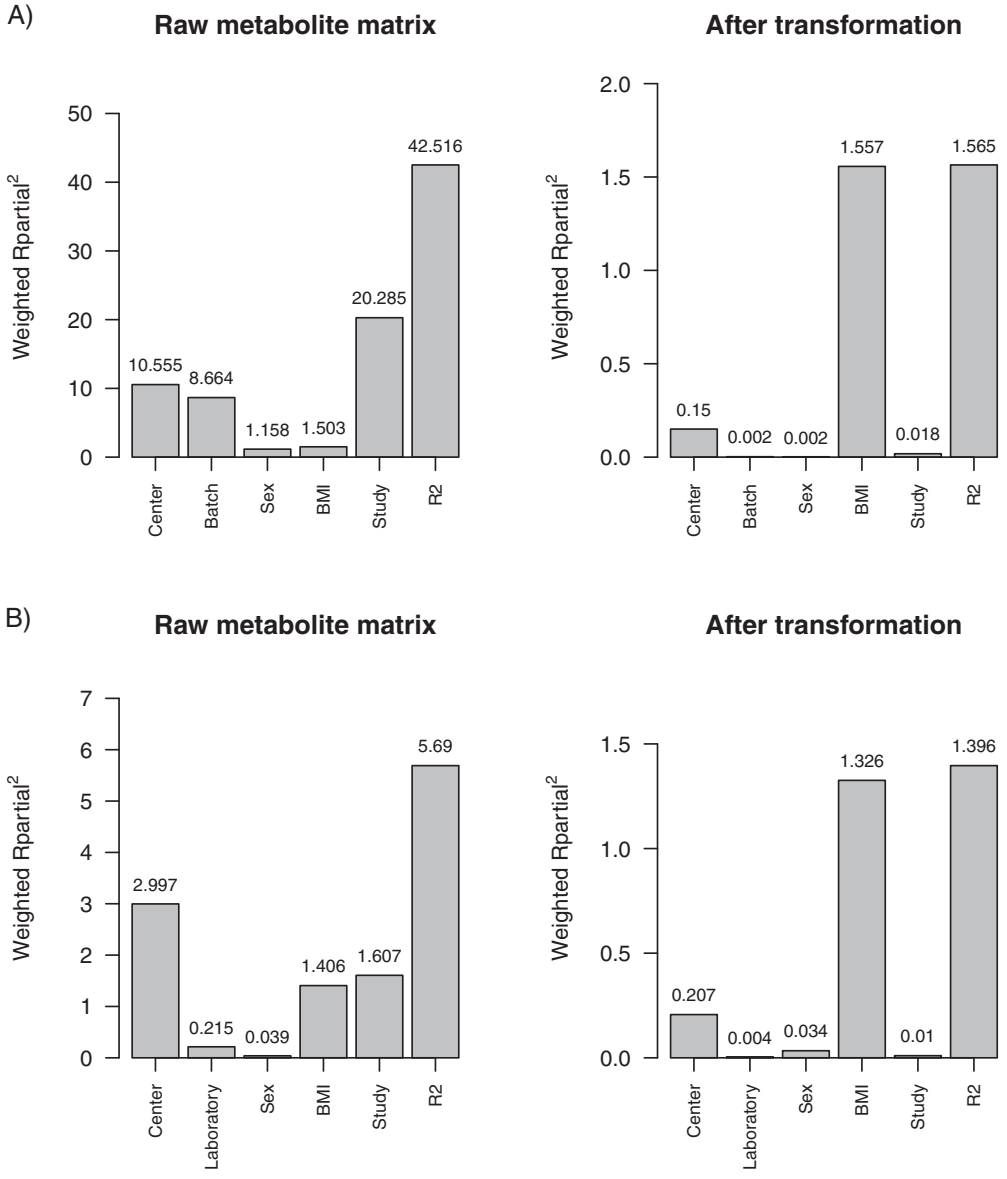
Supplementary References

1. Chajes V, Assi N, Biessy C, et al. A prospective evaluation of plasma phospholipid fatty acids and breast cancer risk in the EPIC study. Ann Oncol 2017;28:2836–2842.
2. Stepien M, Duarte-Salles T, Fedirko V, et al. Alteration of amino acid and biogenic amine metabolism in hepatobiliary cancers: findings from a prospective cohort study. Int J Cancer 2016; 138:348–360.
3. Romisch-Margl W, Prehn C, Bogumil R, et al. Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. Metabolomics 2012; 8:133–142.
4. Schmidt JA, Fensom GK, Rinaldi S, et al. Pre-diagnostic metabolite concentrations and prostate cancer risk in 1077 cases and 1077 matched controls in the European Prospective Investigation into Cancer and Nutrition. BMC Med 2017;15:122.
5. His M, Viallon V, Dossus L, et al. Prospective analysis of circulating metabolites and breast cancer in EPIC. BMC Med 2019;17:178.
6. Assi N, Gunter MJ, Thomas DC, et al. Metabolic signature of healthy lifestyle and its relation with risk of hepatocellular carcinoma in a large European cohort. Am J Clin Nut 2018; 108:117–126.
7. Fages A, Ferrari P, Monni S, et al. Investigating sources of variability in metabolomic data in the EPIC study: the principal component partial R-square (PC-PR2) method. Metabolomics 2014;10:1074–1083.
8. Perrier F, Novoloaca A, Ambatipudi S, et al. Identifying and correcting epigenetics measurements for systematic sources of variation. Clin Epigenetics 2018;10:38.



PC, phosphatidylcholines; SM, sphingomyelins.Concentrations were $\log_2$ transformed.

Supplementary Figure 1. Pearson correlations between 159 endogenous metabolites and 31 fatty acids measured in a subset of 439 colorectal study control participants. Concentrations were $\log_2$ transformed. lysoPC, lysophosphatidylcholine; PC, phosphatidylcholine; SM, sphingomyelins.



Supplementary Figure 2. Variability in discovery metabolomics data explained by different metadata variables as determined by the principal component partial R-square technique. For calculation of metabolic signatures, each column of the metabolite matrix was transformed to the residuals of a mixed-effects model whose explanatory variables were technical confounders: (*A*) 155 endogenous metabolites (n = 1741), and (*B*) 34 fatty acids (n = 4239).

Supplementary Table 1. Summary of WCRF/AICR Recommendations and Scoring System Used in the Present Study

Characteristic	Criteria (operationalization)	Score attributed
Maintain a healthy body weight	BMI, 18.5–24.9	1
	BMI, 25–29.9	0.5
	Other BMI	0
Be moderately physically active, equivalent to brisk walking, for ≥30 min every day	Manual/heavy manual job, or >2 h/wk of vigorous PA, or >30 min/d of cycling/sports	1
	15–30 min/d of cycling or sport	0.5
	<15 min/d of cycling or sport	0
Avoid food and drinks that promote weight gain	Energy dense foods: <125 kcal/100 g/d	1
	125–175 kcal/100 g/d	0.5
	>175 kcal/100 g/d	0
	or sugary drink intake: 0 g/d	1
	0–250 g/d	0.5
	>250 g/d	0
Intake of plant foods	Intake of fruits and vegetables: >400 g/d	1
	200–400 g/d	0.5
	<200 g/d	0
	or dietary fiber intake: >25 g/d	1
	12.5–25 g/d	0.5
	<12.5 g/d	0
Limit intake of animal foods	Intake of red and processed meat or processed meat: <500 g/wk and 3 g/d	1
	<500 g/wk and 3–50 g/d	0.5
	>500 g/wk and >50 g/d	0
Avoid alcohol	Ethanol intake: <20 g/d for men or <10 g/d for women	1
	20–30 g/d for men or 10–20 g/d for women	0.5
	>30 g/d for men or >20 g/d for women	0
Breastfeeding	Cumulative breastfeeding >6 mo	1
	0–6 mo	0.5

BMI, body mass index; PA, physical activity; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.

Supplementary Table 2. Baseline Characteristics for the Discovery Set of EPIC Controls Used to Determine Metabolic Signatures of WCRF/AICR Score

	Participants with endogenous metabolite data	Participants with fatty acid data
N	1741	4239
Study of origin		
Breast	562 (32.3)	2876 (67.8)
Kidney	213 (12.2)	0 (0.0)
Ovary	0 (0.0)	1060 (25.0)
Pancreas	0 (0.0)	303 (7.1)
Prostate	891 (51.2)	0 (0.0)
Liver	75 (4.3)	0 (0.0)
Sex		
Male	1046 (60.1)	118 (2.8)
Female	695 (39.9)	4121 (97.2)
Age at recruitment, *y*	54.50 ± 7.2	53.5 ± 8.1
Height, *cm*	165.6 ± 8.4	161.5 ± 6.8
BMI, kg/m^2	26.8 ± 3.9	25.3 ± 4.2
Total energy intake, *kcal*	2328 ± 670	1964 ± 550
Country		
France	53 (3.0)	638 (15.1)
Italy	903 (51.9)	868 (20.5)
Spain	558 (32.1)	425 (10.0)
United Kingdom	36 (2.1)	825 (19.5)
The Netherlands	11 (0.6)	727 (17.2)
Germany	143 (8.2)	601 (14.2)
Sweden	37 (2.1)	0 (0)
Norway	0 (0)	155 (3.7)
Physical activity, *MET*	81.0 ± 53.9	102.7 ± 53.0
Alcohol intake, *g/d*	18.0 ± 21.5	8.8 ± 12.5
Smoking status		
Nonsmoker	740 (42.5)	2383 (56.2)
Never smoker	564 (32.4)	1046 (24.7)
Smoker	426 (24.5)	729 (17.2)
WCRF/AICR score	2.61 ± 1.01	2.49 ± 1.03
Adherence to individual WCRF/AICR score components (full adherence = 1)		
Weight maintenance	0.56	0.68
Physical activity	0.42	0.40
Intake of foods that promote weight gain	0.59	0.55
Intake of plant foods	0.72	0.60
Intake of animal foods	0.23	0.34
Alcohol intake	0.66	0.79

NOTE. Means and SD or frequency and percentage are shown unless stated otherwise.

BMI, body mass index; EPIC, European Prospective Investigation into Cancer and Nutrition cohort; MET, metabolic equivalent of task; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.

Supplementary Table 3. Details of 155 Endogenous Metabolites and 34 Fatty Acids Measured in Both the Discovery Set and the Colorectal Nested Case–Control Studies

Platform and compound class	Compound name	Exclusion if applicable and reason	Coefficient PLSR model (importance in signature)	CV 1[a]	CV 2[a]
Endogenous metabolites					
Acylcarnitines					
1	C0	Included	-0.017	NA	6.1
2	C10	Missing values		NA	9.2
3	C10:1	Missing values		NA	8.3
4	C12	Missing values		7.4	10.9
5	C12:1	Missing values		7.6	11.8
6	C14	Missing values		8.4	16.5
7	C14:1	Included	0.007	7.2	12.3
8	C14:2	Missing values		10.4	14.1
9	C16	Included	-0.001	8.4	11
10	C16:1	Missing values		12.3	9.4
11	C18	Included	0.006	6.8	15.6
12	C18:1	Included	0.003	7	8.9
13	C18:2	Included	0.001	9.5	10.4
14	C2	Included	-0.011	4.7	6.8
15	C3	Included	0.001	6.1	8.7
16	C4	Included	-0.003	5.5	9.2
17	C5	Included	-0.007	6.8	12
18	C8	Missing values		5	10.5
19	C3-DC (C4-OH)	Missing values		9.3	12.7
20	C4:1	Missing values		10.8	16.5
21	C5-DC (C6-OH)	Missing values		8.8	21
22	C5-M-DC	Missing values		8.6	17.9
23	C7-DC	Missing values		13.2	16.5
24	C9	Missing values		12.8	19.6
25	C5:1-DC	Missing values		12.3	24.1
Amino acids					
26	Alanine	Included	-0.016	6.3	NA
27	Arginine	Included	0.003	5.2	8.1
28	Asparagine	Included	0.02	6.4	NA
29	Aspartate	Included	-0.001	11.5	NA
30	Citrulline	Included	0.013	7.2	NA
31	Glutamine	Included	0.019	7.6	8
32	Glutamate	Included	-0.031	5.7	NA
33	Glycine	Included	0.022	6.9	7.3
34	Histidine	Included	0.005	4.5	7.5
35	Isoleucine	Included	-0.017	7.1	NA
36	Leucine	Included	-0.018	6.9	NA
37	Lysine	Included	-0.008	9.4	NA
38	Methionine	Included	-0.002	11.4	9.5
39	Ornithine	Included	-0.003	11.6	7.2
40	Phenylalanine	Included	-0.011	6.2	8
41	Proline	Included	-0.009	5	6.8
42	Serine	Included	0.023	5	7.3
43	Threonine	Included	0.002	6.1	7.3
44	Tryptophan	Included	0.001	8	7.1
45	Tyrosine	Included	-0.024	6.5	8.3
46	Valine	Included	-0.023	9.1	6.9
Biogenic amines					
47	α-AAA	Missing values		121.2	NA
48	Creatinine	Included	0	3.7	NA
49	Kynurenine	Included	-0.011	7	NA
50	Putrescine	Missing values		35.9	NA
51	Sarcosine	Included	-0.011	8.6	NA
52	Serotonin	Missing values		5.9	NA
53	Spermidine	Missing values		15.5	NA
54	Spermine	Missing values		8.8	NA
55	Transhydroxyproline	Included	-0.019	4.7	NA
56	Taurine	Included	0.001	2.9	NA

Supplementary Table 3. Continued

Platform and compound class	Compound name	Exclusion if applicable and reason	Coefficient PLSR model (importance in signature)	CV 1[a]	CV 2[a]
57	ADMA	Included	0.002	9.3	NA
58	SDMA	Included	0.006	12	NA
LysoPCs					
59	LysoPC a C16:0	Included	-0.003	7.1	6.6
60	LysoPC a C16:1	Included	-0.017	6.7	7.7
61	LysoPC a C17:0	Included	0.035	9	8.3
62	LysoPC a C18:0	Included	0.007	7.5	6.6
63	LysoPC a C18:1	Included	0.02	9.4	6.5
64	LysoPC a C18:2	Included	0.023	8.7	7
65	LysoPC a C20:3	Included	-0.002	7.9	9
66	LysoPC a C20:4	Included	-0.006	9.2	7
67	LysoPC a C28:1	Missing values		12.6	31.7
68	LysoPC a C24:0	Missing values		13.9	14.4
69	LysoPC a C14:0	Missing values		4.7	4.5
70	LysoPC a C28:0	Missing values		20	31.7
Monosaccharides					
71	Hexoses	Included	-0.018	4.9	5.5
PCs, diacyl					
72	PC aa C28:1	Included	-0.004	6.4	8.8
73	PC aa C30:0	Included	-0.015	6.1	9.6
74	PC aa C32:0	Included	-0.01	5.2	7.4
75	PC aa C32:1	Included	-0.037	5.7	10
76	PC aa C32:2	Included	-0.024	8.4	11.5
77	PC aa C32:3	Included	0.003	6.8	9.9
78	PC aa C34:1	Included	-0.019	5.3	7.7
79	PC aa C34:2	Included	-0.009	5.9	6.6
80	PC aa C34:3	Included	-0.016	4.9	7.1
81	PC aa C34:4	Included	-0.031	7.2	7.9
82	PC aa C36:0	Included	0	9.9	11.4
83	PC aa C36:1	Included	-0.015	5.7	7.4
84	PC aa C36:2	Included	-0.008	5.3	6.5
85	PC aa C36:3	Included	-0.012	5.2	6.1
86	PC aa C36:4	Included	-0.033	4.4	5.9
87	PC aa C36:5	Included	-0.011	5.3	9.2
88	PC aa C36:6	Included	-0.005	8.3	13.5
89	PC aa C38:0	Included	0.018	5.1	8.5
90	PC aa C38:3	Included	-0.029	5.1	6.1
91	PC aa C38:4	Included	-0.034	4.9	5.9
92	PC aa C38:5	Included	-0.013	5.4	6.6
93	PC aa C38:6	Included	-0.002	5	8.1
94	PC aa C40:1	Missing values		4.8	13.1
95	PC aa C40:2	Included	0.007	6.7	13.4
96	PC aa C40:3	Included	0.007	11.7	11.3
97	PC aa C40:4	Included	-0.03	4.5	6.4
98	PC aa C40:5	Included	-0.02	6.7	6.5
99	PC aa C40:6	Included	-0.004	8.3	8.2
100	PC aa C42:0	Included	0.011	6.2	9.4
101	PC aa C42:1	Included	0.009	10.5	12.1
102	PC aa C42:2	Included	0.015	6.3	12
103	PC aa C42:4	Included	0.003	7.8	12.3
104	PC aa C42:5	Included	0.004	6.1	11
105	PC aa C42:6	Included	0.004	8	13.8
106	PC aa C24:0	Missing values		36.1	40.3
PCs, acyl-alkyl					
107	PC ae C30:0	Included	0.011	6.1	17.3
108	PC ae C30:2	Included	0.004	13.2	10.2
109	PC ae C32:1	Included	0.005	7.1	9.2
110	PC ae C32:2	Included	0.001	4.6	11.5
111	PC ae C34:0	Included	0.005	7.6	11.2
112	PC ae C34:1	Included	0.015	4.7	7.5
113	PC ae C34:2	Included	0.019	5.2	6.6
114	PC ae C34:3	Included	0.009	4.5	6.7

Supplementary Table 3. Continued

Platform and compound class	Compound name	Exclusion if applicable and reason	Coefficient PLSR model (importance in signature)	CV 1[a]	CV 2[a]
115	PC ae C36:0	Included	-0.009	16.6	13.9
116	PC ae C36:1	Included	0.016	5.8	6.5
117	PC ae C36:2	Included	0.032	5.3	6.6
118	PC ae C36:3	Included	0.015	5.9	6.5
119	PC ae C36:4	Included	-0.02	6	5.9
120	PC ae C36:5	Included	-0.02	4.7	5.7
121	PC ae C38:0	Included	0.007	7	9
122	PC ae C38:2	Included	0.027	10	8.2
123	PC ae C38:3	Included	0.014	7.3	6.8
124	PC ae C38:4	Included	-0.003	5.9	5.8
125	PC ae C38:5	Included	-0.005	6.8	5.8
126	PC ae C38:6	Included	0.001	6.1	6.8
127	PC ae C40:1	Included	0.003	7	12
128	PC ae C40:2	Included	0.01	5.2	8
129	PC ae C40:3	Included	0.022	6.6	7.4
130	PC ae C40:4	Included	0.009	5.5	6.9
131	PC ae C40:5	Included	0.017	5.8	6.2
132	PC ae C40:6	Included	0.032	3.9	7.4
133	PC ae C42:1	Included	0.005	7.4	13.8
134	PC ae C42:2	Included	0.007	6.1	11.6
135	PC ae C42:3	Included	0.014	5.4	10.8
136	PC ae C42:4	Included	0.016	7.7	8.8
137	PC ae C42:5	Included	0.018	6.8	5.6
138	PC ae C44:3	Included	0	13.4	15.7
139	PC ae C44:4	Included	0.015	12	11.3
140	PC ae C44:5	Included	0.015	5.7	7.5
141	PC ae C44:6	Included	0.012	4.5	7.2
Sphingolipids					
142	SM (OH) C14:1	Included	0.014	5.1	7.5
143	SM (OH) C16:1	Included	0.012	8.2	7.1
144	SM (OH) C22:1	Included	0.004	9.9	7.3
145	SM (OH) C22:2	Included	0.015	7.1	7.9
146	SM (OH) C24:1	Included	0.006	12.7	12.5
147	SM C16:0	Included	0.005	8.1	6.5
148	SM C16:1	Included	-0.007	5.2	6.7
149	SM C18:0	Included	-0.016	6.2	6.8
150	SM C18:1	Included	-0.01	5.7	6.5
151	SM C20:2	Included	0.003	23.1	14.7
152	SM C24:0	Included	-0.012	5.9	7
153	SM C24:1	Included	0.001	11.5	7.6
154	SM C26:1	High CV		13.6	25
155	SM C26:0	High CV		17.1	50.6
Fatty acids					
Industrial trans					
1	18:1n-12/9/8t	Included	0.037	13.2	
2	18:2n-6tt	High CV		22.6	
Monounsaturated					
3	14:1n-5	High CV		31.9	
4	15:1	Included	0.049	13.7	
5	16:1n-7/n-9t	Included	0.009	NA	
6	16:1n-7/n-9	Included	-0.058	NA	
7	17:1	Included	0.005	7.3	
8	18:1n-9c	Included	0.041	2.5	
9	18:1n-7c	Included	-0.004	2.1	
10	18:1n-5c	Included	0.029	6.6	
11	20:1n-9c	Included	0.026	2.3	
12	22:1n-9	Included	-0.038	15.6	
13	24:1n-9	Included	-0.035	12.1	
Natural trans					
14	18:1n-7t	High CV		32.7	
15	CLA 9t/11c	Included	-0.016	NA	

Supplementary Table 3. Continued

Platform and compound class	Compound name	Exclusion if applicable and reason	Coefficient PLSR model (importance in signature)	CV 1[a]	CV 2[a]
Polyunsaturated					
16	18:2n-6	Included	0.022	0.7	
17	18:3n-6	Included	0.022	8	
18	20:2n-6c	Included	0.001	1.3	
19	20:3n-9	Included	-0.039	4.1	
20	20:3n-6	Included	0.006	1.4	
21	20:4n-6	Included	-0.011	1.3	
22	22:4n-6	Included	0.014	2.1	
23	22:5n-6	Included	0.039	2.9	
24	18:3n-3	Included	0.029	7.4	
25	20:3n-3	Included	-0.021	8	
26	20:5n-3	Included	-0.02	7	
27	22:5n-3	Included	0.042	1.7	
28	22:6n-3	Included	-0.002	2.7	
Saturated					
29	14:0	Included	0.011	8.6	
30	15:0	Included	0.076	2.7	
31	16:0	Included	-0.043	1.3	
32	17:0	Included	0.149	1.2	
33	18:0	Included	-0.025	1.4	
34	22:0	High CV		28.2	

CV, coefficient of variation; lysoPC, lysophosphatidylcholine; NA, not available; PC, phosphatidylcholine; PLS, partial least-square; QC, quality control; SM, sphingomyelin.

[a]Laboratory 1: International Agency for Research on Cancer; 13 plates of serum samples with 2 QCs per plate for endogenous compounds, 56 batches of plasma samples, 2 QCs per batch for fatty acids. Laboratory 2: Helmholtz Zentrum; 29 plates of serum samples with 5 aliquots of a reference serum as a QC.

Supplementary Table 4. Highest Pearson Correlations Between 159 Endogenous Metabolites and 31 Fatty Acids in 439 Colorectal Study Control Participants

Fatty acid	Endogenous metabolite	Pearson correlation r, $\log_2$ transformed concentrations
PUFA 20:5n-3	PC aa C36:5	0.892
PUFA 22:6n-3	PC aa C38:6	0.767
SFA 14:0	PC aa C30:0	0.746
ITFA 18:1n-12/9/8t	SM C20:2	0.728
PUFA 22:6n-3	PC aa C38:0	0.696
PUFA 22:6n-3	PC aa C40:6	0.694
MUFA 18:1n-9c	PC aa C34:1	0.690
MUFA 16:1n-7/n-9	PC aa C32:1	0.689
PUFA 20:3n-6	PC aa C38:3	0.685
SFA 14:0	PC aa C32:2	0.683
PUFA 20:5n-3	PC aa C36:6	0.669
PUFA 22:4n-6	PC aa C40:4	0.661
PUFA 20:3n-9	PC aa C34:1	0.657
PUFA 20:5n-3	PC ae C38:0	0.653
PUFA 22:6n-3	PC ae C40:6	0.651
PUFA 20:4n-6	PC aa C38:4	0.649
ITFA 18:1n-12/9/8t	PC aa C32:3	0.631
SFA 14:0	PC aa C32:1	0.618
MUFA 18:1n-9c	PC aa C36:1	0.611
SFA 0.625	PC ae C30:0	0.604

ITFA, industrial trans fatty acid; MUFA, monounsaturated fatty acid; PC, phosphatidylcholine; PUFA, polyunsaturated fatty acid; SFA, saturated fatty acid; SM, sphingomyelin.

Supplementary Table 5. Odds Ratios and 95% CI for Individual WCRF/AICR Score Components in the Colorectal Cancer Nested Case–Control Study

Cancer subsite	WCRF/AICR recommendation[a]	OR (95% CI)[b]
Colorectal		
$N = 3216$	Maintain normal body weight	**0.68 (0.67–0.93)**
	Be physically active	**0.87 (0.63–0.99)**
	Limit foods that promote weight gain	**1.10 (0.59–0.99)**
	Eat mostly plant foods	0.93 (0.69–1.26)
	Limit red and processed meat	1.50 (1.13–1.98)
	Avoid alcohol	0.92 (1.77–1.11)
	Overall WCRF score	**0.92 (0.86–1.00)**
Colon		
$N = 2504$	Maintain normal body weight	**0.66 (0.51–0.84)**
	Be physically active	0.85 (0.70–1.04)
	Limit foods that promote weight gain	1.17 (0.77–1.77)
	Eat mostly plant foods	0.91 (0.64–1.28)
	Limit red and processed meat	1.59 (1.17–2.17)
	Avoid alcohol	0.92 (1.74–1.15)
	Overall WCRF score	0.92 (0.84–1.01)
Rectal		
$N = 468$	Maintain normal body weight	0.79 (0.45–1.37)
	Be physically active	0.91 (0.57–1.46)
	Limit foods that promote weight gain	0.65 (0.22–1.89)
	Eat mostly plant foods	0.93 (0.42–2.06)
	Limit red and processed meat	1.10 (1.43–2.83)
	Avoid alcohol	0.78 (0.49–1.22)
	Overall WCRF score	0.89 (0.73–1.09)

NOTE. Boldface indicates statistical significance.
OR, odds ratio; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.
[a]Scored on a scale of 0, 0.5, or 1 according to criteria for individual components.
[b]Adjusted for height, energy intake, highest educational level attained, smoking status, and smoking intensity.

Supplementary Table 6. Additional Sensitivity and Subgroup Analyses in the Nested Case–Control Study

Metabolite platform and anatomic subsite	N	Model[a]	Odds ratio (95% CI) for association per unit increase in the WCRF/AICR score or change in metabolic signature[ab]	
			WCRF/AICR score[a]	Metabolic signature[ab]
Fatty acids				
Colorectal	876	Base co-variates only	**0.78 (0.66–0.91)**	**0.48 (0.28–0.83)**
	876	Base + smoking intensity	**0.77 (0.66–0.91)**	**0.51 (0.29–0.90)**
	876	Base + smoking duration	**0.78 (0.66–0.91)**	**0.49 (0.28–0.85)**
	876	Base + dairy product intake	**0.78 (0.67–0.92)**	**0.50 (0.29–0.88)**
	130	Base + smoking intensity, normal BMI only	–	2.64 (0.25–27.43)
	406	Base + smoking intensity, overweight or obese BMI only	–	0.40 (0.17–0.95)
	210	Base + smoking intensity, WCRF/AICR scores 1 or 2	–	0.38 (0.11–1.33)
	246	Base + smoking intensity, WCRF/AICR scores 3, 4 or 5	–	0.82 (0.23–2.93)
	768	Base model, cases diagnosed after 2 years of follow-up only	**0.84 (0.71–0.99)**	**0.54 (0.30–0.97)**
Endogenous				
Colorectal	3210	Base co-variates only	0.93 (0.85–1.02)	**0.61 (0.49–0.77)**
	3210	Base + smoking intensity	0.93 (0.85–1.02)	**0.62 (0.50–0.78)**
	3210	Base + smoking duration	0.93 (0.85–1.02)	**0.62 (0.49–0.77)**
	3210	Base + dairy product intake	0.94 (0.86–1.03)	**0.62 (0.49–0.77)**
	478	Base + smoking intensity, normal BMI only	–	1.22 (0.63–2.36)
	1352	Base + smoking intensity, overweight or obese BMI only	–	**0.50 (0.35–0.71)**
	722	Base + smoking intensity, WCRF/AICR scores 1 or 2	–	**0.56 (0.35–0.90)**
	848	Base + smoking intensity, WCRF/AICR scores 3, 4 or 5	–	0.69 (0.43–1.11)
	2860	Base model, cases diagnosed after 2 years of follow-up only	0.94 (0.86–1.03)	**0.63 (0.50–0.80)**
Colon	2504	Base co-variates only	0.92 (0.84–1.01)	**0.63 (0.49–0.81)**
	2504	Base + smoking intensity	0.93 (0.85–1.02)	**0.65 (0.50–0.84)**
	2504	Base + smoking duration	0.93 (0.85–1.01)	**0.63 (0.49–0.82)**
	2504	Base + dairy product intake	0.93 (0.85–1.01)	**0.63 (0.49–0.81)**
	2274	Base model, cases diagnosed after 2 years of follow-up only	0.93 (0.85–1.02)	**0.64 (0.49–0.84)**
Rectal	468	Base co-variates only	0.94 (0.78–1.14)	**0.53 (0.31–0.91)**
	468	Base + smoking intensity	0.89 (0.72–1.08)	**0.44 (0.25–0.79)**
	468	Base + smoking duration	0.95 (0.79–1.14)	**0.54 (0.32–0.93)**
	468	Base + dairy product intake	0.97 (0.80–1.17)	**0.55 (0.32–0.95)**
	366	Base model, cases diagnosed after 2 years of follow-up only	0.91 (0.74–1.12)	**0.48 (0.26–0.89)**

NOTE. Boldface indicates statistical significance.
BMI, body mass index; WCRF/AICR, World Cancer Research Fund/American Institute for Cancer Research.
[a]Base models were adjusted for height, energy intake, highest educational level attained, and smoking status.
[b]Measurement of metabolic signature is defined as the metabolite predicted WCRF/AICR score derived from partial least-square regression models fit with endogenous metabolite and fatty acid data in the discovery set.