Identifying Metabolomic Mediators of the Physical Activity and Colorectal Cancer Relationship

Nikos Papadimitriou[1], Nabila Kazmi[2,3], Konstantinos K. Tsilidis[4,5], Rebecca C. Richmond[2,3], Brigid M. Lynch[6,7,8], Benedetta Bendinelli[9], Fulvio Ricceri[10], Maria-Jose Sánchez[11,12,13], Camino Trobajo-Sanmartín[13,14,15], Paula Jakszyn[16,17], Vittorio Simeon[18], Gianluca Severi[19,20], Vittorio Perduca[20,21], Therese Truong[20], Pietro Ferrari[1], Pekka Keski-Rahkonen[1], Elisabete Weiderpass[22], Fabian Eichelmann[23,24], Matthias B. Schulze[23,25], Verena Katzke[26], Renée Turzanski Fortner[26,27], Alicia K. Heath[5], Dagfinn Aune[5,27,28], Rhea Harewood[29], Christina C. Dahm[30], Adrian Llorente[31], Marc J. Gunter[1,5], Neil Murphy[1], and Sarah J. Lewis[2,3]

ABSTRACT

Background: Current evidence suggests higher physical activity (PA) levels are associated with a reduced risk of colorectal cancer. However, the mediating role of the circulating metabolome in this relationship remains unclear.

Methods: Targeted metabolomics data from 6,055 participants in the European Prospective Investigation into Cancer and Nutrition cohort were used to identify metabolites associated with PA and derive a metabolomic signature of PA levels. PA levels were estimated using the validated Cambridge PA index based on baseline questionnaires. Mediation analyses were conducted in a nested case–control study (1,585 cases, 1,585 controls) to examine whether individual metabolites and the metabolomic signature mediated the PA–colorectal cancer association.

Results: PA was inversely associated with colorectal cancer risk (OR per category change: 0.90, 95% confidence interval, 0.83–0.97; P value = 0.009). PA levels were associated with 24 circulating metabolites after FDR correction, with the strongest associations observed for phosphatidylcholine acyl-alkyl (PC ae) C34:3 (FDR-adjusted P value = 1.18×10^{-10}) and lysophosphatidylcholine acyl C18:2 (FDR-adjusted P value = 1.35×10^{-6}). PC ae C34:3 partially mediated the PA–colorectal cancer association (natural indirect effect: 0.991, 95% confidence interval, 0.982–0.999; P value = 0.04), explaining 7.4% of the association. No mediation effects were observed for the remaining metabolites or the overall PA metabolite signature.

Conclusions: PC ae C34:3 mediates part of the PA–colorectal cancer inverse association, but further studies with improved PA measures and extended metabolomic panels are needed.

Impact: These findings provide insights into PA-related biological mechanisms influencing colorectal cancer risk and suggest potential targets for cancer prevention interventions.

[1]Nutrition and Metabolism Branch, International Agency for Research on Cancer, Lyon, France. [2]MRC Integrative Epidemiology Unit (IEU), Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom. [3]Population Health Sciences, The Medical School, University of Bristol, Bristol, United Kingdom. [4]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece. [5]Cancer Epidemiology and Prevention Research Unit, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom. [6]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Australia. [7]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia. [8]Physical Activity Laboratory, Baker Heart and Diabetes Institute, Melbourne, Australia. [9]Clinical Epidemiology Unit, Institute for Cancer Research, Prevention and Cinical Network (ISPRO), Florence, Italy. [10]Department of Clinical and Biological Sciences, Centre for Biostatistics, Epidemiology, and Public Health (C-BEPH), University of Turin, Turin, Italy. [11]Escuela Andaluza de Salud Pública (EASP), Granada, Spain. [12]Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain. [13]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. [14]Instituto de Salud Pública y Laboral de Navarra, Pamplona, Spain. [15]Navarre Institute for Health Research (IdiSNA), Pamplona, Spain. [16]Unit of Nutrition and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of Oncology (ICO-IDIBELL), Barcelona, Spain. [17]Blanquerna School of Health Sciences, Ramon Llull University, Barcelona, Spain. [18]Unit of Medical Statistics, Department of Mental and Physical Health and Preventive Medicine, University of Campania, Naples, Italy. [19]Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Florence, Italy. [20]Paris-Saclay University, UVSQ, Inserm, Gustave Roussy, CESP, Villejuif, France. [21]Université Paris Cité, CNRS, MAP5, Paris, France. [22]International Agency for Research on Cancer, Lyon, France. [23]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany. [24]German Center for Diabetes Research (DZD), Neuherberg, Germany. [25]Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany. [26]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [27]Department of Research, Cancer Registry of Norway, Norwegian Institute of Public Health, Oslo, Norway. [28]Department of Nutrition, Oslo New University College, Oslo, Norway. [29]Department of Surgery and Cancer, Imperial College London, London, United Kingdom. [30]Department of Public Health, Aarhus University, Aarhus Centrum, Denmark. [31]Subdirection of Public Health of Guipuzcoa, San Sebastián, Spain.

Corresponding Author: Sarah J Lewis, Population Health Sciences, The Medical School, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK. E-mail: S.J.Lewis@bristol.ac.uk

Cancer Epidemiol Biomarkers Prev 2025;34:578–87

doi: 10.1158/1055-9965.EPI-24-1390

Introduction

Colorectal cancer is one of the most common cancer types globally, responsible for almost 2 million new cancer cases and over 900,000 related deaths in 2020 (1). Prospective cohort studies and Mendelian randomization analyses have consistently found that higher levels of physical activity (PA) are associated with reduced risk of colorectal cancer and especially colon cancer (2–4). More specifically, World Cancer Research Fund (WCRF) in its last report ranked the evidence regarding the protective role of PA on colon cancer risk as convincing, whereas no conclusion was drawn for rectal cancer (2). Similarly, an umbrella review of the literature concluded that the protective association of recreational PA with colon cancer was supported by strong evidence (3). A recent Mendelian randomization study further reported that an increase in accelerometer-measured PA levels was associated with lower risk of overall colorectal cancer and colon cancer but not with rectal cancer (4). Potential mechanisms underlying this relationship are incompletely understood but are believed to include a reduction in insulin resistance and inflammation through direct effects of exercise as well as the lowering of body weight (2). However, it is possible that currently unidentified, biological pathways mediate the PA and colorectal cancer relationship.

Metabolomics involves the systematic identification and quantification of multiple metabolites in an organism or biological sample to explore associations with disease (5). The metabolome directly reflects the physiologic and pathologic state of an individual, and these measurements can be used to provide insights into potential mechanistic pathways involved in carcinogenesis (6–8).

Several prior studies, both cross-sectional and prospective, have investigated associations between different levels of PA and metabolite levels (9, 10). PA-related metabolomic alterations identified have included changes in fatty acid metabolism, mobilization and lipolysis, the tricarboxylic acid cycle, glycolysis, amino acid metabolism, carnitine metabolism, purine metabolism, cholesterol metabolism and insulin sensitivity (10). Although there have been some studies investigating the role of metabolites associated with PA and other cancer types such as breast cancer (11), there have been no similar studies for colorectal cancer.

In the current analysis, we used targeted blood metabolomics data from the European Prospective Investigation into Cancer and Nutrition (EPIC) study to identify individual metabolites associated with PA and derive a metabolomic signature of PA levels by combining data from all identified metabolites. We then examined the extent to which the identified metabolomic signature and individual metabolites mediated the inverse association observed between PA levels and colorectal cancer risk in a case–control study nested in EPIC.

Materials and Methods

EPIC

EPIC is a multicenter cohort of 521,330 participants from the general populations of 10 European countries (Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the UK) who were recruited between 1991 and 2000 (12, 13). The participants were followed up for an average of 15.2 years. The current study used data from all EPIC countries apart from Greece, Sweden, and Norway for which biosamples and/or data were unavailable due to administrative constraints.

Outcome assessment

Incident cancer cases were identified using cancer registries in UK, Spain, Italy (except the center in Naples), the Netherlands, and Denmark. For France, Germany, and Naples, incident cancer cases were identified during follow-up from a combination of sources including cancer and pathology centers, health insurance records, and active follow-up of study participants. Colorectal cancer cases were defined using the 10th revision of the International Classification of Diseases and the third revision of the International Classification of Diseases for Oncology codes C18–C20 corresponding to cancers occurring between the cecum and rectum.

Exposure assessment

PA levels were estimated using questionnaire data collected at baseline, focusing on past-year PA in occupational, leisure, and household domains, and classified according to the validated Cambridge PA index into four categories: inactive, moderately inactive, moderately active, and active (14). The index was based on four EPIC questions on habitual PA during the past year (Supplementary Methods S1) and was validated in 173 participants from a continuing population-based cohort study in Ely (15), Cambridgeshire, where four measures of cardiorespiratory fitness and four measures of 4-day energy expenditure by heart-rate monitoring were completed across 1 year. At the final visit, participants completed the PA questionnaire that refers to activity in the past 12 months (14). Its repeatability was further assessed in a sample of 2,271 participants from the EPIC–Norfolk cohort (weighted $\kappa = 0.6$; $P < 0.0001$; ref. 14). More details regarding what each level of PA index represents can be found in the Supplementary Table S1.

Confounder assessment

At baseline, questionnaires were used to collect information on demographics, medical conditions, and behavioral factors, including smoking status, alcohol consumption, level of education, and diabetes status. Anthropometric measurements were collected using standardized methods (13). Validated country- and center-specific dietary questionnaires were used to collect information on diet.

Blood collection and laboratory analysis

Blood samples were collected at baseline following a standardized protocol in France, Germany, Italy, the Netherlands, Spain, and the UK (13). Serum, plasma, erythrocytes, and buffy coat aliquots were stored in liquid nitrogen (−196°C) in a centralized biobank at the International Agency for Research on Cancer (IARC). In Denmark, blood fractions were stored locally in the vapor phase of liquid nitrogen containers (−150°C). Sample analyses were performed at the IARC and the Helmholtz Zentrum, München, Germany, using the targeted AbsoluteIDQ p150 or p180 Kits (BIOCRATES Life Sciences AG) to measure concentrations of 171 metabolites in serum or plasma (depending on sample's availability; ref. 16). Samples were assayed on different LC and MS instruments across the different studies, but each study used one single pair of LC-MS instruments for all samples (17). A previous analysis examining data quality for the Biocrates kit showed that for typical biological samples (serum and plasma from healthy individuals) the median interlaboratory coefficient of variation was 7.6%, with 85% of metabolites exhibiting a median interlaboratory coefficient of variation of <20%, demonstrating the reproducibility of the method (18). Additionally, previous targeted metabolomics studies in different EPIC centers further confirmed that by reporting equally low

coefficients of variation (19, 20) or high intraclass correlation coefficients (21).

Data normalization

To derive the metabolomic signature of PA levels, a recently developed analytical pipeline was applied to a dataset of 15,428 participants from seven distinct (breast, colorectal, endometrial, gallbladder, kidney, liver, and prostate cancer) case–control studies nested within EPIC with available metabolomics data from prediagnostic blood samples (17). This pipeline involved three steps to pool together and normalize the participant data across the six separate nested case–control studies to increase the statistical power while correcting for several preanalytical and analytical factors that might induce artificial differences (17): Step (i) exclusions of the least informative observations and metabolites and imputation of missing data; Step (ii) identification of the main components of variability; Step (iii) application of linear mixed models to remove unwanted variability. Metabolites and samples with >20% missing data and samples assayed in batches with less than 10 samples were excluded (17 samples and 54 metabolites). For the remaining metabolites, measurements below the limit of detection or lower limit of quantification were set to half the batch-specific limit of detection or lower limit of quantification, respectively. Metabolite measurements above the upper limit of quantification were set to values equal to the upper limit of quantification. Fully missing values were set to the batch-specific median of non-missing values if less than 50% of the measurements in the batch were missing and to the study-specific median of the batch specific medians otherwise (17). The current study included 117 metabolites (15 acylcarnitines, 13 amino acids, 76 glycerophospholipids, 1 monosaccharide, and 12 sphingolipids). The principal component partial R-square method was used to calculate the total variation in the metabolomics data and its source (22). Based on the principal component partial R-square results, the main sources of variation were preanalytical and analytical variables, including the individual cancer nested case–control study, batch effects, and EPIC center, which collectively explained 40% of the total variation of the metabolomics measurements (Supplementary Fig. S1A). The final step included the normalization of the log-transformed metabolite data by implementing a linear mixed model to correct for variation due to these factors, which were included as random effects in the model (Supplementary Fig. S1B).

Statistical analysis

Metabolites of PA

Adjusted linear regression models to investigate the associations between the normalized metabolite data (dependent variables) with the levels of PA were applied to the controls only after excluding the participants from the colorectal cancer nested case–control study (n = 6,055). The PA index was entered into the model as a continuous variable, and all the metabolite-associated coefficients corresponded to a category change in the PA levels. The model was adjusted for several *a priori* defined covariates, including sex, age at blood collection (continuous), fasting status at blood collection (<3, 3–6, >6 hours, or unknown), education (none, primary school, technical/professional school, secondary school, longer education including university degree, or unknown/unspecified), smoking status at recruitment (current, former, never, or unknown), body mass index (BMI; continuous, kg/m^2), alcohol consumption (continuous, g/day), diabetes status (yes, no, or unknown), and daily intakes of total energy, red and processed meat, and fruits and vegetables (quartiles). The FDR was calculated to correct for multiple comparisons (FDR-adjusted P values <0.05 denoted statistical significance, corresponding to a nominal P value of 0.01). We also calculated the effects of PA on the identified metabolites among the controls in the colorectal cancer dataset to examine their concordance.

Metabolomic signature of PA

The metabolomic signature of the PA index for the metabolites associated with PA in the previous step was derived through a partial least squares (PLS) regression (23). Briefly, PLS extracted linear combinations, referred to as PLS factors, of the PA index and the identified metabolites, allowing a simultaneous decomposition of both sets of variables with the aim of maximizing their covariance. The metabolomic signature was then correlated with the PA index by applying Spearman correlation coefficient.

Metabolomic signature of PA and colorectal cancer risk

Using the PLS factors derived in the previous step, the metabolomic signature was then derived and investigated in relation to colorectal cancer risk in a new dataset of 1,585 colorectal cancer cases and 1,585 controls from a case–control study nested within EPIC in which cases and controls were matched by recruitment center, sex, age at blood collection (±3 to ±5 years in subsequent rounds), date (±1 to ±6 months) and time of day of blood sampling (±1 to ±3 hours), and fasting status. Multivariable conditional logistic regression models were used to assess the associations between the metabolomic signature of PA, as well as individual metabolites, and colorectal cancer risk. The models were adjusted for the same variables as those used to identify the metabolomic signature of PA.

Mediating role of metabolites on PA and colorectal cancer risk

We investigated the possible mediating role of the PA metabolomic signature and each metabolite separately on the association between PA and colorectal cancer. Estimates of the natural direct effect (NDE), the natural indirect effect (NIE), and the total effect (TE) of PA index were calculated using a counterfactual approach adapted to dichotomous outcomes (24). The term NDE refers to the effect of PA index that is not mediated by the potential mediating metabolites, and the NIE is that part of the effect of PA index that is mediated. Formulae from VanderWeele and Vansteelandt were adapted to accommodate continuous exposures and use of conditional logistic regression models (25). In summary, two main models were applied to obtain NDE, NIE, and the mediator effect of metabolite adjusted for the PA effect. In the outcome model, PA and the mediator variables (one at a time) were both included as independent variables in a conditional logistic regression for colorectal cancer risk. In models for each mediator of interest, the mediator was linearly regressed on PA only on the subset of controls to account for the nested case–control design (26). The TE was obtained from a conditional logistic regression of PA with colorectal cancer risk. Assuming that the outcome was rare, we calculated the proportion mediated that captures the importance of the mediating pathway and is defined on the risk difference scale, and it is calculated based on the following formula $\frac{NDE^{*}(\mathrm{NIE}-1)}{(NDE^{*}\mathrm{NIE}-1)}$ (26) which ranges from 0% to 100%, and it is meaningful when NDE and NIE have the same direction of association. All models were adjusted for the same potential confounders as listed above. Further details on how the estimates were obtained can be found on previously published work within EPIC (27, 28). After additional testing, no

exposure–mediator interactions were observed and therefore were not included in the final mediation analyses. We also conducted the mediation analysis including colon cancer cases only as existing literature suggests that the PA inverse association is stronger for colon cancer than for rectal cancer (2).

Sensitivity analyses

We repeated our mediation analyses excluding cancer cases that were diagnosed within 2 and 5 years of follow-up following blood draw (lag analysis) to account for potential bias due to underlying subclinical cancer.

A previous analysis in EPIC showed that obesity, measured by waist circumference, mediated the inverse association between high PA levels and colon cancer risk (29), indicating that obesity may be on the causal pathway between PA and colon cancer risk rather than a confounder. Therefore, we reran the models to calculate the metabolomic signature of PA and its mediating effect on the association between PA and colorectal cancer without BMI adjustment to investigate the potential impact of adjusting for the BMI in our initial results.

All analyses were conducted in R (version 4.3.3, http://www.r-project.org). The R package plsgenomics was used to conduct the PLS regression, and the pls.regression.cv function was applied to determine the best number of latent components to be used for PLS regression. A flowchart recapping the different steps of the analysis and a directed acyclic graph for the mediation analysis are presented in **Fig. 1** and Supplementary Fig. S2, respectively.

All participants provided written informed consent to participate in the EPIC study. This study was approved by the ethics committee of the IARC and all centers. This study complies with the Declaration of Helsinki.

Data availability

Data from EPIC are not publicly available, but access requests can be made to the to the EPIC Steering Committee (https://epic.iarc.fr/access/). The code to perform the metabolite normalization is available in the relevant publication (17).

Results

Table 1 shows the baseline characteristics of the study participants. For participants in the metabolomic signature discovery phase, the average age at blood collection was 57 years of age, 56% of the participants were men, and 48% of the controls were controls in a study of prostate cancer. Overall, 59% of participants were categorized as inactive or moderately inactive, 62% of participants were classified as overweight or obese (**Table 1**). Comparison of the cancer cases with the controls in the colorectal cancer study revealed that the controls tended to be more physically active, smoked less, were less obese, and consumed more fruits and vegetables than colorectal cancer cases (**Table 1**).

Metabolomic signature of PA

After FDR correction, a higher level of PA was associated with increased circulating levels of 14 metabolites (all glycerophospholipids) and decreased levels of 10 metabolites (3 sphingolipids, 2 acylcarnitines, 2 amino acids, 2 glycerophospholipids, and sum of hexoses; **Fig. 2**; Supplementary Table S2). The most significant associations were observed for phosphatidylcholine acyl-alkyl (PC ae) C34:3 (FDR-adjusted P value $= 1.18 \times 10^{-10}$) and lysophosphatidylcholine acyl (lysoPC a) C18:2 (FDR-adjusted P value $= 1.35 \times 10^{-6}$; both positively associated with the PA index). Out of the 24 identified metabolites, 19 showed concordant effect estimates in our smaller colorectal cancer control dataset (Supplementary Table S3).

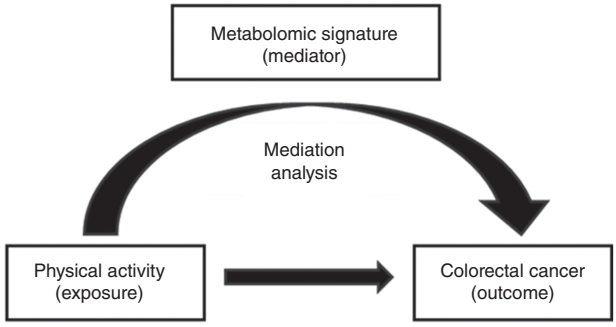

Figure 1.
A directed acyclic graph for the mediation analysis. The direct arrow from the exposure (PA) to the outcome (colorectal cancer) corresponds to the NDE of PA that is not mediated by the potential mediating metabolites. The second arrow that connects PA to colorectal cancer through the metabolomic signature corresponds to the NIE that is part of the effect of PA that is mediated by the metabolomic signature.

Based on the PLS regression and the pls.regression.cv function, the metabolic signature of PA was derived as the loadings on the first latent variable. The metabolites with the largest contribution to the signature were lysoPC a C18:2 (loading = 0.40), lysoPC a C18:1 (loading = 0.38), PC ae C34:3 (loading = 0.33), and PC ae C36:3 (loading = 0.31; Supplementary Fig. S3). There were some correlations among the metabolites, mostly clustered around three main classes of metabolites (phosphatidylcholines, lysophosphatidylcholines, and sphingomyelins; Supplementary Fig. S4). The metabolomic signature of PA was correlated with the PA index (Spearman index = 0.18; P value $< 2.2 \times 10^{-16}$).

Examining the mediating role of metabolites in the PA and colorectal cancer association

Overall, PA was inversely associated with colorectal cancer risk [TE OR per category change: 0.90; 95% confidence interval (CI), 0.83–0.97; P value = 0.009; **Table 2**)]. After FDR correction among the 24 identified metabolites, none was significantly associated with colorectal cancer risk, although three metabolites; C3 (OR per 1-SD change: 0.69; 95% CI, 0.53–0.89; P value = 0.005), PC ae C34:3 (OR per 1-SD change: 0.66; 95% CI, 0.50–0.88; P value = 0.005), and PC ae C36:3 (OR per 1-SD change: 0.70; 95% CI, 0.49–0.99; P value = 0.04) showed nominally significant associations. The NDE estimates for the direct association of PA with colorectal cancer risk were almost identical to the TE estimates. There was some evidence of mediation for PC ae C34:3 (NIE: 0.991; 95% CI, 0.982–0.999; P value = 0.04) suggestive of a mediating effect of 7.4% (**Table 2**). For the overall metabolomic signature of PA, the NIE showed no to weak mediation effect (NIE: 0.994; 95% CI, 0.987–1.002; P value = 0.21; **Table 2**).

A similar pattern of results was observed for colon cancer, with an inverse association observed with higher levels of PA and colon cancer risk (TE OR per category change: 0.92, 95% CI, 0.84–1.00; P value = 0.05). The results from the mediation analysis were in general similar to the overall colorectal cancer analysis

Table 1. Sociodemographic, lifestyle, dietary, and blood-sampling related characteristics of participants included in the analyses.

	EPIC		
	Metabolomic signature discovery phase (*n* = 6,055)	**Colorectal cancer nested case–control dataset (*n* = 1,585 pairs)**	
		Cases	Control
Age, mean (SD)	57 (7.9)	57 (7.5)	57 (7.5)
Sex, *n* (%)			
Men	3,367 (56)	868 (55)	868 (55)
Women	2,688 (44)	717 (45)	717 (45)
Educational level, *n* (%)			
None	377 (6)	140 (9)	127 (8)
Primary school	2,021 (33)	548 (35)	603 (39)
Technical/professional	1,397 (23)	351 (22)	352 (22)
Secondary	770 (13)	246 (15)	205 (13)
University or higher	1,280 (21)	250 (16)	259 (16)
Not specified	210 (4)	210 (4)	39 (2)
BMI (kg/m^2), mean (SD)	26 (3.9)	27 (4.4)	26 (3.9)
Categories of BMI, *n* (%)			
<18.5	39 (1)	10 (1)	15 (1)
18.5–25	2,267 (37)	539 (34)	586 (37)
25–30	2,786 (46)	700 (44)	727 (46)
30+	963 (16)	336 (21)	257 (16)
PA, *n* (%)			
Inactive	1,445 (24)	464 (29)	406 (25)
Moderately inactive	2,145 (35)	542 (34)	551 (35)
Moderately active	1,337 (22)	309 (20)	312 (20)
Active	1,128 (19)	270 (17)	316 (20)
Smoking status, *n* (%)			
Never	2,632 (43)	679 (43)	751 (47)
Former	2,029 (34)	515 (32)	474 (30)
Smoker	1,353 (22)	385 (24)	352 (22)
Unknown	41 (1)	6 (1)	8 (1)
Red and processed meat intake (g/day), mean (SD)	84 (55)	85 (64)	86 (52)
Fruits and vegetables intake (g/day), mean (SD)	443 (262)	452 (268)	473 (272)
Energy intake (kcal), mean (SD)	2,258 (647)	2,212 (633)	2,229 (631)
Alcohol intake (g/day), mean (SD)	16 (20)	17 (21)	15 (19)
Substudy controls, *n* (%)			
Breast	1,516 (25)		
Endometrial	765 (12)		
Gallbladder	47 (1)		
Kidney	587 (10)		
Liver	221 (4)		
Prostate	2,919 (48)		
Participants by country, *n* (%)			
France	197 (3)	52 (3)	52 (3)
Italy	1,262 (21)	391 (25)	391 (25)
Spain	1,105 (18)	317 (20)	317 (20)
United Kingdom	1,156 (19)	228 (14)	228 (14)
the Netherlands	490 (8)	126 (8)	126 (8)
Germany	1,582 (26)	164 (10)	164 (10)
Denmark	263 (4)	307 (19)	307 (19)

(Supplementary Table S4). PC ae C34:3 again had the strongest proportion of effect mediated at 7.4%, as in the overall colorectal cancer analysis, but the evidence for the NIE estimate was weak (NIE: 0.994; 95% CI, 0.987–1.001; *P* value = 0.10). There was little evidence of mediation for the overall metabolomic signature of PA (NIE: 0.995; 95% CI, 0.987–1.005; *P* value = 0.34).

The 2-year lag analysis showed an attenuated mediating effect of 4.1% for PC ae C34:3 (NIE: 0.994; 95% CI, 0.987–1.001; *P* value = 0.09) which was further reduced to 2.8% (NIE: 0.996; 95% CI, 0.989–1.002; *P* value = 0.23) in the 5-year lag analysis (Supplementary Table S4). Again, little evidence of mediation for the overall metabolomic signature was found regardless of the lag period (Supplementary Table S4).

Finally, after repeating our analysis without BMI adjustment, 30 metabolites were associated with PA, but most of them were the same as in our initial analysis (Supplementary Table S5). The mediation analysis without BMI adjustment moreover showed some mediating effects for the overall metabolomic signature, as well as

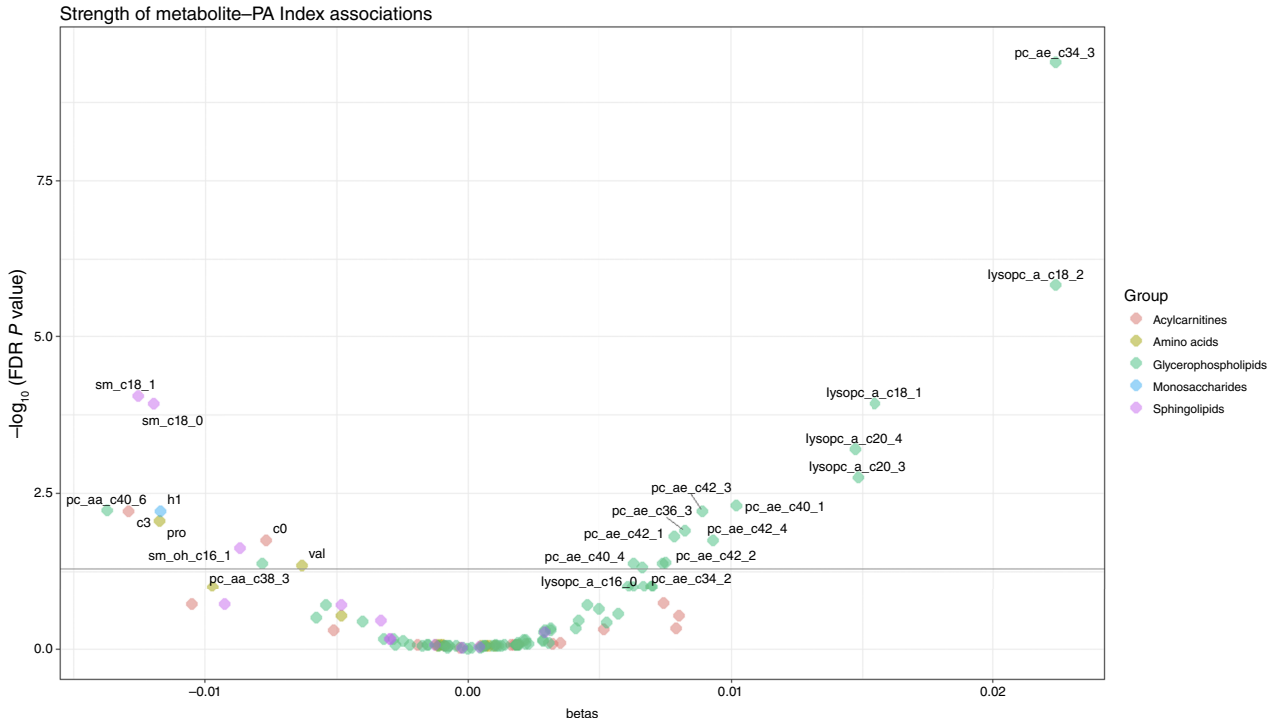

Figure 2.

Volcano plot with associations between metabolites with levels of PA as measured by PA index. The metabolites above the horizontal line showed a significant association with PA after correcting for multiple comparisons.

for PC ae C34:3. PC ae C34:3 mediated 9.7% (NIE: 0.988; 95% CI, 0.979–0.997; *P* value = 0.01), whereas the proportion mediated for the metabolomic signature was 10.5% (NIE: 0.987; 95% CI, 0.978–0.997; *P* value = 0.02; Supplementary Table S6). Similar results were observed in the analysis of colon cancer (proportions mediated: 9.3% and 13.4%, respectively), whereas the estimates were attenuated in the lag analyses (Supplementary Table S6).

Discussion

In this analysis, we identified a novel circulating metabolomic signature of PA that comprised several phosphatidylcholines and lysophosphatidylcholines. Overall, we found little evidence of any mediating role of the overall PA-related metabolomic signature in the inverse association of PA with colorectal cancer risk. However, among the individual PA-related metabolites, we found some evidence of PC ae C34:3 as a possible mediator, although this effect was attenuated when we introduced a lag time into our analysis.

Prior metabolomic studies for PA have focused mostly on the effect of different training regimes on the metabolome (10, 30). Currently, there are relatively few population-based cohort studies focusing on habitual PA (31–34). A targeted metabolomics analysis of 5,197 participants in Nurses' Health Study I and II and the Health Professionals Follow-up Study identified 20 metabolites significantly associated with higher levels of habitual PA measured as metabolomic equivalent of task-hours per week based on leisure time activities (31), with most of the identified metabolites showing positive associations with PA including several phosphatidylcholines and lysophosphatidylcholines including lysoPCs a C18:1 and a C18:2 that were also identified in our study. However, PC ae C34:3 was not in the panel of metabolites investigated in that study (31). An untargeted metabolomics study of 7,271 Finnish men without diabetes at baseline from the Metabolic Syndrome in Men cohort identified 198 metabolites associated with greater leisure time PA (active vs. non-active; ref. 32). Again, PC ae C34:3 was not reported in the results, but in general inverse associations were observed for most phosphatidylcholines and PA, which is not in agreement with our analysis. However, the Finnish study applied the one-way ANOVA test to assess whether the levels of the metabolites differ by PA without any further adjustment for potential confounders, and additionally only men were included in that study, which taken together could have an impact on the results (32). A targeted metabolomics analysis in EPIC-Potsdam (N = 100 participants) reported that PC ae C34:3 was positively associated with cardiorespiratory fitness (0.06, 95% CI, 0.01–0.12; mL/kg/minutes), although the association was slightly attenuated when the models were further adjusted for PA energy expenditure, time spent sedentarily, and time spent in engaging in vigorous activity (0.05, 95% CI, −0.01 to 0.12; mL/kg/minutes; ref. 33). Additionally, a PC-derived factor of 19 metabolites mostly consisting of acyl-alkyl-phosphatidylcholines including PC ae C34:3 was also positively associated with cardiorespiratory fitness (0.074, 95% CI, 0.01–0.14; *P* value = 0.03; ref. 33). Finally, a more recent serum metabolite network analysis in 2,380 participants, again in EPIC-Potsdam, reported that PC ae C34:3 along with other acyl-alkyl-phosphatidylcholines comprised a group of metabolites that was positively associated with cardiorespiratory fitness (34).

Table 2. Results from the mediation analyses, with ORs and their associated 95% CIs for the NDE, NIE, TE, and proportion mediated.

	NDE (95% CI)	NIE (95% CI)	TE (95% CI)	% mediated
PA			0.90 (0.83–0.97)	
Metabolites				
Glyceroph lysoPC a C16:0	0.90 (0.83–0.98)	0.999 (0.995–1.002)		0.9
Glyceroph lysoPC a C18:1	0.90 (0.84–0.98)	0.997 (0.991–1.002)		2.0
Glyceroph lysoPC a C18:2	0.91 (0.84–0.98)	0.994 (0.986–1.002)		3.2
Glyceroph lysoPC a C20:3	0.90 (0.83–0.97)	1.000 (0.997–1.002)		0.4
Glyceroph lysoPC a C20:4	0.90 (0.83–0.97)	1.001 (0.996–1.005)		0
Glyceroph PC aa C38:3	0.90 (0.84–0.98)	0.999 (0.995–1.004)		0.6
Glyceroph PC aa C40:6	0.90 (0.83–0.97)	1.000 (0.999–1.002)		0
Glyceroph PC ae C34:2	0.90 (0.83–0.98)	0.998 (0.994–1.002)		1.7
Glyceroph PC ae C34:3	0.91 (0.84–0.98)	0.991 (0.982–0.999)		7.4
Glyceroph PC ae C36:3	0.90 (0.83–0.98)	0.999 (0.995–1.004)		0.6
Glyceroph PC ae C40:1	0.90 (0.84–0.98)	0.997 (0.992–1.002)		1.8
Glyceroph PC ae C40:4	0.90 (0.83–0.97)	1.001 (0.998–1.005)		0
Glyceroph PC ae C42:1	0.90 (0.83–0.97)	1.000 (0.997–1.002)		0.2
Glyceroph PC ae C42:2	0.90 (0.83–0.97)	1.000 (0.998–1.002)		0
Glyceroph PC ae C42:3	0.90 (0.83–0.98)	0.998 (0.995–1.002)		1.0
Glyceroph PC ae C42:4	0.90 (0.83–0.97)	1.002 (0.998–1.007)		0
Sphingo Sm C18:0	0.90 (0.83–0.97)	1.000 (0.999–1.002)		0
Sphingo Sm C18:1	0.90 (0.83–0.97)	1.002 (0.998–1.002)		0
Sphingo Sm Oh C16:1	0.90 (0.83–0.97)	0.999 (0.995–1.002)		1.0
Acylcarn C0	0.90 (0.83–0.97)	1.002 (0.998–1.006)		0
Acylcarn C3	0.90 (0.83–0.97)	1.007 (1.000–1.015)		0
Amino acid proline	0.90 (0.83–0.97)	0.999 (0.996–1.003)		0.6
Amino acid valine	0.90 (0.83–0.97)	1.001 (0.998–1.003)		0
H1 (overall hexose)	0.90 (0.83–0.98)	0.997 (0.993–1.002)		1.8
Overall signature	0.91 (0.84–0.98)	0.994 (0.987–1.002)	0.96 (0.92–1.01)	4.1

Abbreviations: Acylcarn, acylcarnitine; Sphingo Sm Oh, sphingolipid sphingomyelin (OH).

Previous metabolomics studies have also linked PC ae C34:3 with obesity-related traits (35–37). A previous study in EPIC reported that higher body size was associated with lower levels of PC ae C34:3 (35). Additionally, PC ae C34:3 levels were also lower in participants with metabolic syndrome in the KORA F4 (beta = −0.76, $P = 4.06 \times 10^{-82}$) and SHIP-TREND-0 (beta = −0.73, $P = 1.08 \times 10^{-20}$) studies (N = 3,803 participants, 31% with metabolic syndrome in the combined dataset; ref. 36). Another targeted metabolomics study in EPIC-Potsdam (N = 2,282) found that increased levels of PC ae C34:3 and other acyl-alkyl-phosphatidylcholines were associated with lower risk of type 2 diabetes (37). Finally, data from the Tübingen Family study further reported that PC ae C34:3 was positively correlated with insulin sensitivity (spearman correlation = 0.15) and negatively with insulin secretion (spearman correlation = −0.24; ref. 37). Overall, evidence suggests that circulating concentrations of PC ae C34:3 and other acyl-alkyl-phosphatidylcholines are higher with greater levels of PA and lower with adiposity and in those with type 2 diabetes and metabolic syndrome.

The possible PA-related mediating role of PC ae C34:3 was driven by the inverse association between this metabolite and colorectal cancer risk that was observed in the colorectal cancer nested case–control study. Apart from our study in EPIC, there is currently lack of evidence regarding the role of PC ae C34:3 on colorectal cancer risk. However, a small untargeted metabolomics study of 50 colorectal cancer–control pairs found that the concentration levels of PC O-34:3, a potential isobar of PC ae C34:3, were lower in patients with colorectal cancer than in controls, which is in agreement with our results (38).

Currently, there is a lack of evidence regarding the pathways through which PC ae C34:3 could be associated with decreased colorectal cancer risk, although several studies have consistently reported inverse associations between phosphatidylcholines including PC ae C34:3 and other cancer sites (6, 39–41). It has been suggested that potential mechanisms could be their anti-inflammatory role in the colorectum as well as their antioxidant effects (42, 43). Chronic inflammation is a hallmark of cancer (44) and an important driver of colorectal cancer development (45), while oxidative stress has also been linked with an increased risk of colorectal cancer (46). On the other hand, it has been shown increased levels of PA reduce chronic inflammation and increase adaptive response to oxidative stress (47, 48). We could therefore postulate that PC ae C34:3 could be a link mediating at least partially the effects of PA on colorectal cancer risk.

Removing the BMI from the models lead to increased evidence of mediating effect of the metabolomic signature in our analysis. A previous study in EPIC has found that waist circumference and not the BMI has a mediating role in the association between PA and colon cancer (29). Although the BMI and waist circumference are highly correlated, they represent different obesity phenotypes in which waist circumference represents better abdominal obesity that is closely associated with insulin resistance. Therefore, PA has been suggested to reduce waist circumference and induce insulin sensitivity, independent of changes in body weight (29). Our findings support the hypothesis that general obesity could lie on the pathway linking PA to colorectal cancer or that obesity is a crude biomarker of the metabolic changes that occur in an obesogenic environment that lead to colorectal cancer and therefore caution is needed when adjusting for the BMI.

The main strength of our study is the large sample size due to the pooling of multiple metabolomic datasets from six substudies within EPIC. We used a novel recently developed analytical pipeline to

increase statistical power while correcting for several preanalytical and analytical factors that might have induced artificial differences in the data (17). Also, no participants from the colorectal cancer nested case–control study were used in the metabolite identification phase, thus avoiding sample overlap that could bias our results. Nevertheless, our study has also some limitations. Information on PA levels was assessed once from baseline questionnaires that are prone to measurement error due to misreporting of the PA levels. The conduct of a clinical trial perhaps would be a more efficient study design to investigate the metabolomic signature of PA, given the greater control over the definition and measurement of PA; however, using data from a large scale study like EPIC allows us to gain a better insight into the longer-term metabolomic effects (10). A validation study within EPIC has reported only a moderate correlation between the Cambridge PA index and objectively measured PA (Pearson correlation index with PA energy expenditure: $r^2 = 0.33$; 95% CI, 0.28–0.38; ref. 49). Therefore, we know that the PA index is an imperfect measure of PA. However, of the three indices evaluated in the validation study (total PA index and recreational index are the other two), the Cambridge PA index showed the stronger associations with objectively measured PA (49). Under nondifferential measurement error with a normally distributed mediator, the NIE will be biased toward the null, and if direct and indirect effects are in the same direction, the bias of the NDE is away from the null (50). In the current analysis, this may have led to an underestimation of the indirect effects, and an overestimation of the direct effects, resulting in a lower mediated proportion. Also, even though we adjusted for several confounders, these may not have been captured well in the measurement, and residual confounding could still occur. Additionally, Biocrates' lipid metabolites are not specific, so one name could match several lipids that can have different biological roles. Our study was also limited by the relatively small number of measured metabolites. In addition, we cannot exclude the possibility that larger sample size was needed to identify mediating effects of a smaller magnitude. Moreover, 19 out of the 24 metabolites that were identified at the first stage of our analysis showed concordant results in the smaller colorectal cancer dataset; however, additional replications in external datasets would be required in order to assess the robustness of the associations. Finally, a crucial key assumption of our approach was the temporal ordering between exposure, mediator, and outcome. In our study, the PA index that estimates participants' PA levels over the 12 months preceding enrolment was assessed at baseline at the same time as blood sample collection. Temporality was better ensured in our analyses with incident colorectal cancer owing to the prospective study design. However, we cannot exclude potential reverse causality in our results, especially given the attenuation of the mediating effect observed in the lag analysis.

We identified a novel metabolomic signature of PA levels but found little evidence of this signature mediating the associations between PA and colorectal cancer except in the analysis that was not adjusted for the BMI. We found novel evidence of one metabolite, PC ae C34:3, as having a possible mediating role in the PA and colorectal cancer relationship. However, additional high-quality epidemiologic and experimental studies are needed to validate our results.

Authors' Disclosures

R.T. Fortner reports grants from the German Federal Ministry of Education and Research that supported EPIC-Heidelberg during the conduct of the study. No disclosures were reported by the other authors.

Disclaimer

Where authors are identified as personnel of the IARC/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 IARC/World Health Organization. This article is the result of the scientific work of N. Murphy while he was affiliated at IARC.

Authors' Contributions

N. Papadimitriou: Formal analysis, writing–original draft, writing–review and editing. **N. Kazmi:** Writing–review and editing. **K.K. Tsilidis:** Writing–review and editing. **R.C. Richmond:** Writing–review and editing. **B.M. Lynch:** Writing–review and editing. **B. Bendinelli:** Writing–review and editing. **F. Ricceri:** Writing–review and editing. **M.-J. Sánchez:** Writing–review and editing. **C. Trobajo-Sanmartín:** Writing–review and editing. **P. Jakszyn:** Writing–review and editing. **V. Simeon:** Writing–review and editing. **G. Severi:** Writing–review and editing. **V. Perduca:** Writing–review and editing. **T. Truong:** Writing–review and editing. **P. Ferrari:** Writing–review and editing. **P. Keski-Rahkonen:** Writing–review and editing. **E. Weiderpass:** Writing–review and editing. **F. Eichelmann:** Writing–review and editing. **M.B. Schulze:** Writing–review and editing. **V. Katzke:** Writing–review and editing. **R.T. Fortner:** Writing–review and editing. **A.K. Heath:** Writing–review and editing. **D. Aune:** Writing–review and editing. **R. Harewood:** Writing–review and editing. **C.C. Dahm:** Writing–review and editing. **A. Llorente:** Writing–review and editing. **M.J. Gunter:** Supervision, methodology, writing–review and editing. **N. Murphy:** Supervision, methodology, writing–review and editing. **S.J. Lewis:** Conceptualization, supervision, methodology, writing–review and editing.

Acknowledgments

Funding for grant IIG_FULL_2020_019 was obtained from Wereld Kanker Onderzoek Fonds (WKOF) as part of the World Cancer Research Fund International grant programme awarded to Professor S.J. Lewis. In addition, K.K. Tsilidis, R.C. Richmond, M.J. Gunter, N Murphy, and S.J. Lewis were also supported by Cancer Research UK (grant number C18281/A29019) awarded to Profs R. Martin and C. Relton. The authors thank all EPIC centers for provision of biological samples and participant data. The coordination of EPIC is financially supported by the International Agency for Research on Cancer (IARC) and by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, with infrastructure support provided by the NIHR Imperial Biomedical Research Centre (BRC). The national cohorts are supported by Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM; France); German Cancer Aid, German Cancer Research Center (DKFZ), 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, Compagnia di SanPaolo and National Research Council (Italy); The National Institute for Public Health and the Environment (RIVM), Bilthoven, Dutch Ministry of Public Health, Welfare, and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (the Netherlands); 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); Cancer Research UK (14136 to EPIC-Norfolk; C8221/A29017 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk; MR/M012190/1 to EPIC-Oxford; UK). IDIBELL acknowledges support from the Generalitat de Catalunya through the CERCA Program. The colorectal cancer studies were funded by the World Cancer Research Fund (reference: 2013/1002; www.wcrf.org/) and the European Commission (FP7: BBMRI-LPC; reference: 313010; https://ec.europa.eu/). For the purpose of open access, the author has applied a 'Creative Commons Attribution (CC BY) public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

Note

Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Received September 24, 2024; revised November 27, 2024; accepted January 28, 2025; published first January 30, 2025.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209–49.
2. World Cancer Research Fund/American Institute for Cancer Research. Continuous update project expert report 2018. diet, nutrition, physical activity and colorectal cancer. [cited 2025 Jan 25]. Available from: https://www.wcrf.org/wp-content/uploads/2024/10/Colorectal-cancer-report.pdf.
3. Rezende LFM, Sá THd, Markozannes G, Rey-López JP, Lee I.M, Tsilidis K.K, et al. Physical activity and cancer: an umbrella review of the literature including 22 major anatomical sites and 770 000 cancer cases. Br J Sports Med 2018;52:826–33.
4. Papadimitriou N, Dimou N, Tsilidis KK, Banbury B, Martin RM, Lewis SJ, et al. Physical activity and risks of breast and colorectal cancer: a Mendelian randomisation analysis. Nat Commun 2020;11:597.
5. Wang W, Rong Z, Wang G, Hou Y, Yang F, Qiu M. Cancer metabolites: promising biomarkers for cancer liquid biopsy. Biomark Res 2023;11:66.
6. His M, Viallon V, Dossus L, Gicquiau A, Achaintre D, Scalbert A, et al. Prospective analysis of circulating metabolites and breast cancer in EPIC. BMC Med 2019;17:178.
7. McCullough ML, Hodge RA, Campbell PT, Stevens VL, Wang Y. Prediagnostic circulating metabolites and colorectal cancer risk in the cancer prevention study-II nutrition cohort. Metabolites 2021;11:156.
8. Mrowiec K, Kurczyk A, Jelonek K, Debik J, Giskeødegård GF, Bathen T.F, et al. Association of serum metabolome profile with the risk of breast cancer in participants of the HUNT2 study. Front Oncol 2023;13:1116806.
9. Sakaguchi CA, Nieman DC, Signini EF, Abreu RM, Catai AM. Metabolomics-based studies assessing exercise-induced alterations of the human metabolome: a systematic review. Metabolites 2019;9:164.
10. Kelly RS, Kelly MP, Kelly P. Metabolomics, physical activity, exercise and health: a review of the current evidence. Biochim Biophys Acta Mol Basis Dis 2020;1866:165936.
11. Watts EL, Moore SC, Abar L, Hong HG, Saint-Maurice PF, O'Connell C, et al. Physical activity, metabolites, and breast cancer associations. J Natl Cancer Inst 2024:djae246.
12. Slimani N, Kaaks R, Ferrari P, Casagrande C, Clavel-Chapelon F, Lotze G, et al. European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study: rationale, design and population characteristics. Public Health Nutr 2002;5:1125–45.
13. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr 2002;5:1113–24.
14. Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, et al. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Public Health Nutr 2003;6:407–13.
15. Williams DR, Wareham NJ, Brown DC, Byrne CD, Clark PM, Cox BD, et al. Undiagnosed glucose intolerance in the community: the isle of ely diabetes project. Diabet Med 1995;12:30–5.
16. Rothwell JA, Murphy N, Bešević J, Kliemann N, Jenab M, Ferrari P, et al. Metabolic signatures of healthy lifestyle patterns and colorectal cancer risk in a European cohort. Clin Gastroenterol Hepatol 2022;20:e1061.
17. Viallon V, His M, Rinaldi S, Breeur M, Gicquiau A, Hemon B, et al. A new pipeline for the normalization and pooling of metabolomics data. Metabolites 2021;11:631.
18. Siskos AP, Jain P, Römisch-Margl W, Bennett M, Achaintre D, Asad Y, et al. Interlaboratory reproducibility of a targeted metabolomics platform for analysis of human serum and plasma. Anal Chem 2017;89:656–65.
19. Floegel A, von Ruesten A, Drogan D, Schulze MB, Prehn C, Adamski J, et al. Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-potsdam. Eur J Clin Nutr 2013;67:1100–8.
20. Carayol M, Leitzmann MF, Ferrari P, Zamora-Ros R, Achaintre D, Stepien M, et al. Blood metabolic signatures of body mass index: a targeted metabolomics study in the EPIC cohort. J Proteome Res 2017;16:3137–46.
21. Carayol M, Licaj I, Achaintre D, Sacerdote C, Vineis P, Key T.J, et al. Reliability of serum metabolites over a two-year period: a targeted metabolomic approach in fasting and non-fasting samples from EPIC. PLoS One 2015;10:e0135437.
22. Fages A, Ferrari P, Monni S, Dossus L, Floegel A, Mode N, et al. Investigating sources of variability in metabolomic data in the EPIC study: the Principal Component Partial R-square (PC-PR2) method. Metabolomics 2014;10:1074–83.
23. Tenenhaus M. La régression PLS: Théorie et pratique. Paris: Editions Technip; 1998.
24. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 2013;18:137–50.
25. Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. Am J Epidemiol 2010;172:1339–48.
26. Vanderweele TJ, Explanation in causal inference: Methods for mediation and interaction. New York (NY): Oxford University Press; 2015. p. 728.
27. Assi N, Rinaldi S, Viallon V, Dashti SG, Dossus L, Fournier A, et al. Mediation analysis of the alcohol-postmenopausal breast cancer relationship by sex hormones in the EPIC cohort. Int J Cancer 2020;146:759–68.
28. Assi N, Thomas DC, Leitzmann M, Stepien M, Chajès V, Philip T, et al. Are metabolic signatures mediating the relationship between lifestyle factors and hepatocellular carcinoma risk? Results from a nested case-control study in EPIC. Cancer Epidemiol Biomarkers Prev 2018;27:531–40.
29. Aleksandrova K, Jenab M, Leitzmann M, Bueno-de-Mesquita B, Kaaks R, Trichopoulou A, et al. Physical activity, mediating factors and risk of colon cancer: insights into adiposity and circulating biomarkers from the EPIC cohort. Int J Epidemiol 2017;46:1823–35.
30. Jaguri A, Al Thani AA, Elrayess MA. Exercise metabolome: insights for health and performance. Metabolites 2023;13:694.
31. Ding M, Zeleznik OA, Guasch-Ferre M, Hu J, Lasky-Su J, Lee IM, et al. Metabolome-wide association study of the relationship between habitual physical activity and plasma metabolite levels. Am J Epidemiol 2019;188:1932–43.
32. Kemppainen SM, Fernandes Silva L, Lankinen MA, Schwab U, Laakso M. Metabolite signature of physical activity and the risk of type 2 diabetes in 7271 men. Metabolites 2022;12:69.
33. Wientzek A, Floegel A, Knüppel S, Vigl M, Drogan D, Adamski J, et al. Serum metabolites related to cardiorespiratory fitness, physical activity energy expenditure, sedentary time and vigorous activity. Int J Sport Nutr Exerc Metab 2014;24:215–26.
34. Floegel A, Wientzek A, Bachlechner U, Jacobs S, Drogan D, Prehn C, et al. Linking diet, physical activity, cardiorespiratory fitness and obesity to serum metabolite networks: findings from a population-based study. Int J Obes (Lond) 2014;38:1388–96.
35. Kliemann N, Viallon V, Murphy N, Beeken RJ, Rothwell JA, Rinaldi S, et al. Metabolic signatures of greater body size and their associations with risk of colorectal and endometrial cancers in the European Prospective Investigation into Cancer and Nutrition. BMC Med 2021;19:101.
36. Shi M, Han S, Klier K, Fobo G, Montrone C, Yu S, et al. Identification of candidate metabolite biomarkers for metabolic syndrome and its five components in population-based human cohorts. Cardiovasc Diabetol 2023;22:141.
37. Floegel A, Stefan N, Yu Z, Mühlenbruch K, Drogan D, Joost HG, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. Diabetes 2013;62:639–48.
38. Fernandes A, Messias MCF, Duarte GHB, de Santis GKD, Mecatti GC, Porcari AM, et al. Plasma lipid profile reveals plasmalogens as potential biomarkers for colon cancer screening. Metabolites 2020;10:262.
39. Breeur M, Ferrari P, Dossus L, Jenab M, Johansson M, Rinaldi S, et al. Pan-cancer analysis of pre-diagnostic blood metabolite concentrations in the European Prospective Investigation into Cancer and Nutrition. BMC Med 2022;20:351.
40. Schmidt JA, Fensom GK, Rinaldi S, Scalbert A, Appleby PN, Achaintre D, et al. Patterns in metabolite profile are associated with risk of more aggressive prostate cancer: a prospective study of 3,057 matched case-control sets from EPIC. Int J Cancer 2020;146:720–30.
41. Guida F, Tan VY, Corbin LJ, Smith-Byrne K, Alcala K, Langenberg C, et al. The blood metabolome of incident kidney cancer: a case-control study nested within the MetKid consortium. PLoS Med 2021;18:e1003786.
42. Treede I, Braun A, Sparla R, Kühnel M, Giese T, Turner JR, et al. Anti-inflammatory effects of phosphatidylcholine. J Biol Chem 2007;282:27155–64.
43. Wallner S, Schmitz G. Plasmalogens the neglected regulatory and scavenging lipid species. Chem Phys Lipids 2011;164:573–89.

44. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov 2022;12:31–46.
45. Schmitt M, Greten FR. The inflammatory pathogenesis of colorectal cancer. Nat Rev Immunol 2021;21:653–67.
46. Dash C, Bostick RM, Goodman M, Flanders WD, Patel R, Shah R, et al. Oxidative balance scores and risk of incident colorectal cancer in a US Prospective Cohort Study. Am J Epidemiol 2015;181:584–94.
47. Ji LL, Gomez-Cabrera MC, Vina J. Exercise and hormesis: activation of cellular antioxidant signaling pathway. Ann N Y Acad Sci 2006;1067:425–35.
48. Gleeson M, Bishop NC, Stensel DJ, Lindley MR, Mastana SS, Nimmo MA. The anti-inflammatory effects of exercise: mechanisms and implications for the prevention and treatment of disease. Nat Rev Immunol 2011;11:607–15.
49. The InterAct Consortium; Peters T, Brage S, Westgate K, Franks PW, Gradmark A, et al. Validity of a short questionnaire to assess physical activity in 10 European countries. Eur J Epidemiol 2012;27:15–25.
50. VanderWeele TJ, Valeri L, Ogburn EL. The role of measurement error and misclassification in mediation analysis: mediation and measurement error. Epidemiology 2012;23:561–4.