His et al. BMC Medicine (2021) 19:312
https://doi.org/10.1186/s12916-021-02183-2

BMC Medicine

RESEARCH ARTICLE Open Access

# Lifestyle correlates of eight breast cancer-related metabolites: a cross-sectional study within the EPIC cohort

Mathilde His[1], Vivian Viallon[1], Laure Dossus[1], Julie A. Schmidt[2], Ruth C. Travis[2], Marc J. Gunter[1], Kim Overvad[3], Cecilie Kyrø[4], Anne Tjønneland[4,5], Lucie Lécuyer[6], Joseph A. Rothwell[6], Gianluca Severi[6,7], Theron Johnson[8], Verena Katzke[8], Matthias B. Schulze[9,10], Giovanna Masala[11], Sabina Sieri[12], Salvatore Panico[13], Rosario Tumino[14], Alessandra Macciotta[15], Jolanda M. A. Boer[16], Evelyn M. Monninkhof[17], Karina Standahl Olsen[18], Therese H. Nøst[18], Torkjel M. Sandanger[18], Antonio Agudo[19,20], Maria-Jose Sánchez[21,22,23,24], Pilar Amiano[25,26,27], Sandra M. Colorado-Yohar[27,28,29], Eva Ardanaz[27,30,31], Linda Vidman[32], Anna Winkvist[33], Alicia K. Heath[34], Elisabete Weiderpass[35], Inge Huybrechts[1] and Sabina Rinaldi[1*]

## Abstract

**Background:** Metabolomics is a promising molecular tool for identifying novel etiological pathways leading to cancer. In an earlier prospective study among pre- and postmenopausal women not using exogenous hormones, we observed a higher risk of breast cancer associated with higher blood concentrations of one metabolite (acetylcarnitine) and a lower risk associated with higher blood concentrations of seven others (arginine, asparagine, phosphatidylcholines (PCs) aa C36:3, ae C34:2, ae C36:2, ae C36:3, and ae C38:2).

**Methods:** To identify determinants of these breast cancer-related metabolites, we conducted a cross-sectional analysis to identify their lifestyle and anthropometric correlates in 2358 women, who were previously included as controls in case-control studies nested within the European Prospective Investigation into Cancer and Nutrition cohort and not using exogenous hormones at blood collection. Associations of each metabolite concentration with 42 variables were assessed using linear regression models in a discovery set of 1572 participants. Significant associations were evaluated in a validation set ($n$ = 786).

**Results:** For the metabolites previously associated with a lower risk of breast cancer, concentrations of PCs ae C34:2, C36:2, C36:3, and C38:2 were negatively associated with adiposity and positively associated with total and saturated fat intakes. PC ae C36:2 was also negatively associated with alcohol consumption and positively associated with two scores reflecting adherence to a healthy lifestyle. Asparagine concentration was negatively associated with adiposity. Arginine and PC aa C36:3 concentrations were not associated to any of the factors examined. For the metabolite previously associated with a higher risk of breast cancer, acetylcarnitine, a positive association with age was observed.

* Correspondence: rinaldis@iarc.fr
[1]International Agency for Research on Cancer (IARC/WHO), Nutrition and Metabolism Branch, 150 cours Albert Thomas, 69372, CEDEX 08 Lyon, France
Full list of author information is available at the end of the article

© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Conclusions:** These associations may indicate possible mechanisms underlying associations between lifestyle and anthropometric factors, and risk of breast cancer. Further research is needed to identify potential non-lifestyle correlates of the metabolites investigated.

**Keywords:** Metabolites, Breast cancer, Cross-sectional, Lifestyle, Anthropometry

## Background

Metabolomics is an important tool in the identification of new etiological pathways associated with chronic diseases, including breast cancer [1–8], as the metabolome reflects both endogenous parameters and exogenous exposures [9]. Prospective studies using targeted metabolomics (analyses of a pre-defined panel of metabolites) or untargeted metabolomics approaches have reported novel associations of pre-diagnostic blood concentrations of endogenous metabolites with breast cancer risk. These metabolites include lysophosphatidylcholine a C18:0 [8], 16a-hydroxy-DHEA-3-sulfate [4, 5], various carnitines [4, 5], caprate (10:0) [6], histidine, glycerol, N-acetyl-glycoprotein [7], acetone, glycerol-derived compounds, other amino acids, and lipids [2, 3], suggesting new potential avenues of research and possible additional targets for prevention.

In a previous case-control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, we investigated the association between blood concentrations of endogenous metabolites, measured by targeted metabolomics, and risk of breast cancer [1]. We reported a positive association between acetylcarnitine (C2) and breast cancer risk and negative associations of arginine, asparagine, phosphatidylcholines acyl-alkyl (PCs ae) C36:3, C34:2, C36:2, C38:2, and phosphatidylcholine diacyl (PC aa) C36:3 with breast cancer risk, among women not using exogenous hormones at blood collection.

To further assess how these findings can inform breast cancer prevention research, a better understanding of potentially modifiable determinants of blood levels of these metabolites is needed. Towards this aim, we report here the results of a cross-sectional analysis nested in the EPIC cohort to investigate associations of a wide range of lifestyle and anthropometric variables and acetylcarnitine, arginine, asparagine, PCs aa C36:3, ae C34:2, ae C36:2, ae C36:3, and ae C38:2.

## Methods

### The EPIC study

EPIC is an ongoing multi-center cohort study including approximately 520,000 participants recruited between 1992 and 2000 from ten European countries [10]. Female participants ($n$ = 367,903) were aged 35–75 years at recruitment. Detailed information was collected on dietary, lifestyle, reproductive, medical, and anthropometric data at inclusion [10]. Around 246,000 women from all countries provided a baseline blood sample. Blood was collected according to a standardized protocol in France, Germany, Greece, Italy, the Netherlands, Norway, Spain, and the UK [10]. Serum (except in Norway), plasma, erythrocytes, and buffy coat aliquots were stored in liquid nitrogen (−196°C) in a centralized biobank at IARC. In Denmark, blood fractions were stored locally in the vapor phase of liquid nitrogen containers (−150°C), and in Sweden, they were stored locally at −80°C in standard freezers. All participants provided written informed consent to participate in the EPIC study. This study was approved by the ethics committee of the International Agency for Research on Cancer (IARC) and all centers.

### Study population and cross-sectional design

This study included all female EPIC participants (1) who provided a blood sample; (2) who were previously included in one of six case-control studies on cancer etiology nested within the EPIC cohort (on breast [1], endometrial [11], colorectal [12], kidney [13], liver [14], and gallbladder cancers) with available blood concentrations of acetylcarnitine, arginine, asparagine, PCs aa C36:3, ae C34:2, ae C36:2, ae C36:3, and ae C38:2 measured by the same targeted metabolomics approach; (3) who were included as control participants in these studies (i.e., free of cancer (except non-melanoma skin cancer) at the time of the diagnosis of the cases, using incidence-density sampling, and matched to cases by age, sex, study center, time of blood collection, fasting status at blood collection (except for kidney cancer study), menopausal status and exogenous hormone use at blood collection (for breast, endometrial, liver, and gallbladder studies), and phase of menstrual cycle (for breast and endometrial cancer studies)); and (4) whose samples were included in an analytical batch including at least 10 samples, to ensure proper normalization of metabolite concentrations (see the "Statistical analyses" section) ($N$ = 3163).

We then excluded women who declared use of hormones at blood collection ($n$ = 768), and those whose hormone use status at blood collection was unknown ($n$ = 37), because associations between the studied metabolites and breast cancer risk were limited to hormone non-users [1]. The current analysis included data from 2358 participants.

The 2358 participants were split into a discovery set ($N$ = 1572, 66.7% of the population) and a validation set ($N$ = 786, 33.3% of the population). Metabolites of interest were those found to be associated with breast cancer risk, and this observed association could result from associations between metabolites and some of the correlates under study in the present work. Thus, the discovery set included all controls from the breast cancer study ($n$ = 1079), and randomly selected controls from the other nested case-control studies ($n$ = 493), while the validation set did not include participants from the breast cancer study. This way, associations identified on the discovery set and further validated on the validation set are guaranteed not to be driven by the breast cancer study only.

### Laboratory measurements

Before exclusions of hormone users, a total of 3179 samples were available for 3163 women. All samples, plasma (in 95.1% of samples) or serum, were assayed by liquid chromatography-mass spectrometry using the AbsoluteIDQ p180 commercial kit (Biocrates Life Sciences AG, Innsbruck, Austria). A total of 2289 (72.0%) samples were assayed at the laboratory of the Biomarkers Group at IARC (breast, colorectal, kidney, and liver cancer studies); 851 (26.8%) at the Imperial College, London; and 39 (1.2%) at the Helmholtz Zentrum, München, Germany. At IARC, analyses were run on a QTRAP5500 (breast, kidney, and liver cancer studies) and TQ4500 (colorectal cancer study) mass spectrometers (AB Sciex, Framingham, MA, USA), while at the Imperial College London and Helmholtz Zentrum, analyses were run using an API4000TQ (endometrial and gallbladder cancer studies). All analyses for a given study were performed using the same instrument. Sixteen participants had their samples analyzed in two different studies, at IARC and at the Helmholtz Zentrum, for whom the metabolite concentrations were averaged over the two measures.

Out of the 3179 samples, arginine concentrations could not be quantified in five, as they were below the lower limit of quantification (LLOQ) and were therefore imputed to half this LLOQ, consistently with previous work [1].

### Covariate data

Details of data collection in EPIC are described elsewhere [10]. Lifestyle and medical factors were assessed in the baseline questionnaire. Usual dietary intakes were assessed using center- or country-specific validated questionnaires covering the previous 12 months and matched to the US Department of Agriculture food composition database to estimate macronutrient intakes [15]. Glycemic index and glycemic load were computed. In all EPIC centers, except France, Oxford, and Norway, height, weight, and waist and hip circumference were measured on all participants using similar protocols (in Umeå (Sweden), only weight and height were measured). In France and Oxford, weight, height, and waist and hip circumferences were measured in a sub-set of participants, but self-reported weight and height were obtained from all individuals, and validation studies showed high correlations between self-reported and measured values ($r \geq 0.90$) [16, 17]. In Oxford, self-reported measurements also included waist and hip circumferences. In Norway, only self-reported height and weight were available.

Dietary data were used to compute the inflammatory score of the diet (ISD) [18] (reflecting the inflammatory potential of the diet based on 28 dietary components), the modified Mediterranean diet score [19] (a 9-component score indicating the degree of adherence to the traditional Mediterranean diet; 0 minimal adherence to 9 maximal adherence), and the Diet Quality Index-International (DQI-I; a 17-component score based on general nutritional guidelines [20, 21]; 0 to 100, minimal to maximal diet quality). Dietary and lifestyle data were combined to calculate the Healthy Lifestyle Index (HLI) [22], designed to reflect five components of lifestyle factors (smoking, alcohol consumption, diet (cereal fibers, red and processed meat, the ratio of polyunsaturated to saturated fatty acids, margarine, glycemic load, and fruits and vegetables), physical activity, and body mass index; ranging from 0, least healthy, to 20). Furthermore, we calculated the World Cancer Research Fund/American Institute for Cancer Research score, which reflects recommendations for cancer prevention on weight maintenance, physical activity, intake of food and drinks which promote weight gain, of plant-based foods, of animal-based foods, of alcohol, and breastfeeding [23] (from 0, low adherence to recommendation, to 7 for women).

### Statistical analyses

#### *Normalization of metabolite concentrations*

A specific statistical pipeline was developed [24] and applied on raw metabolite concentrations (before exclusion of hormone users) to adequately pool measures obtained from different studies, instruments, and laboratories. This pipeline was shown to be efficient in removing unwanted variability and improving the comparability of measurements acquired across different nested studies. Log-transformed concentrations of the metabolites of interest were normalized to remove effects of analytical batch and study, which were estimated as random effects in mixed-effects linear models correcting for possible heteroscedasticity. Corrected metabolite concentrations

analyzed in this work correspond to residuals from the model.

### Missing data

When missing values on covariates represented less than 5% of the overall values, they were imputed to the mode value (categorical variables: number of full-term pregnancies, ever use of oral contraceptive, ever use of hormones for menopause (by menopausal status), education level, physical activity, smoking status, fasting status) or median (continuous variables: age at menarche, age at first full-term pregnancy (among parous women), duration of breastfeeding among women who breastfed, waist circumference, hip circumference, waist/hip ratio, time at blood collection). When missing values represented more than 5% of values for a variable, this variable was categorized, and a "missing" category was created (phase of menstrual cycle at blood collection for pre- and perimenopausal women, breastfeeding, lifetime alcohol consumption, Healthy Lifestyle Index, WCRF/AICR score).

### Identification of correlates

Participants' characteristics were described using frequencies for categorical variables and mean (standard deviation) for continuous variables. We calculated partial Pearson's correlations between metabolite concentrations (adjusted for center and age) and between metabolites and age (adjusted for center).

Analyses were first run in the discovery set. For each metabolite of interest and each lifestyle variable, a linear regression model was built with metabolite concentration as a dependent variable. Models were adjusted for center of recruitment, age at blood collection, menopausal status (premenopausal, perimenopausal, postmenopausal [25]), phase of the menstrual cycle for premenopausal women (follicular, ovulatory, luteal, missing), time of the day, and fasting status at blood collection ("No": < 3 h since last meal (< 4 h in Umeå), "In between": 3–6 h (4–8 h in Umeå), and "Yes": > 6 h (> 8 h in Umeå)). Models that examined age as exposure were not adjusted for age, and models with menopausal status as main exposure were not adjusted for phase of menstrual cycle, as this variable is defined in premenopausal women only.

Variables tested as possible correlates were age at blood collection (continuous), age at menarche (continuous), total duration of menstrual cycles (quartiles/missing), pregnancy (ever/never), number of full-term pregnancies (continuous), age at first full-term pregnancy (nulliparous/quartiles), breastfeeding (ever/never/missing), duration of breastfeeding (nulliparous/quartiles/missing), use of oral contraceptive (ever/never; current users excluded), menopausal status at blood collection (premenopausal/perimenopausal/postmenopausal), use of hormones for menopause (ever/never; current users are excluded), education level (no schooling or primary/technical, professional or secondary/longer education), physical activity (Cambridge Index [26]: inactive/moderately inactive/moderately active/active), smoking status (never/former/current), smoking status combined with intensity (never/current, 1–15 cigarettes/day/current, 16+ cigarettes/day/current, pipe/cigar/occasional/former, quit for ≤10 years/former, quit 11–20 years/former, quit > 20 years), baseline alcohol consumption (continuous, g/day), lifetime alcohol consumption (non-drinker/former drinker/current > 0–3 g/day/> 3–12 g/day/> 12–24 g/day/> 24 g/day/missing), BMI (continuous, $kg/m^2$), waist circumference (continuous, cm), hip circumference (continuous, cm), waist/hip ratio (continuous), height (continuous, cm), total energy intake (continuous, kcal/day), and the following food components estimated as residuals on total energy intake (continuous, g/day): protein, carbohydrate, starch, sugar, fiber, fat (total), fatty acids (monounsaturated, polyunsaturated, saturated, trans, trans-monoenoic, trans-polyenoic), glycemic index (continuous), glycemic load (continuous), Healthy Lifestyle Index (0–10/11–15/16–20), WCRF/AICR score (quartiles/missing), modified Mediterranean diet score (continuous), diet quality index (continuous), and inflammatory score of the diet (continuous).

For each metabolite, *P*-values from *F*-tests for each variable were collected and were corrected for multiple testing by controlling for family-wise error rate at $\alpha = 0.05$ by permutation-based stepdown *minP* adjustment of *P*-values, a method which accounts for dependencies between tests [27].

### Validation

All statistically significant associations in the discovery set (based on *P*-values corrected for multiple tests ≤0.05) were assessed in the validation set, using the same model and categories of variables as in the discovery set. In this validation set, a more conservative approach was chosen for controlling for multiple tests [28], i.e., the Bonferroni correction based on the number of tests run for each metabolite.

For all variables showing a significant association with the metabolites of interest in both the discovery and validation sets, continuous variables were categorized (quartiles) and means of metabolites, with 95% confidence intervals, were estimated in each category, using the overall dataset ($n = 2358$).

### Interactions

For each metabolite and each variable examined as potential correlate, we investigated interaction with fasting

status (no/in between/yes), menopausal status at blood collection (pre-/peri-/postmenopausal), and BMI (18.5–24.9/25–29.9/≥30 kg/m$^2$, excluding $n$ = 15 participants with BMI < 18.5 kg/m$^2$), in the discovery set. To do so, an interaction term was added in the model and the *P*-value associated with this term was evaluated, after correction for multiple testing using the permutation *minP* algorithm.

#### *Sensitivity analyses*

We conducted sensitivity analyses (1) excluding participants from the liver and gallbladder studies ($n$ = 128), for which the blood fraction analyzed was serum and not plasma, and (2) excluding participants with self-reported diabetes ($n$ = 71) or with missing data on diabetes status ($n$ = 160) at recruitment.

## Results

Participants' characteristics overall and from the discovery and validation sets are shown in Table 1. Overall, 39.7% of the participants were not fasting at blood collection while 44.4% were considered fasting (more than 6 h since last meal (8 h in Umeå)). Around 30% of participants were premenopausal. Overall, participant characteristics were similar among discovery and validation sets (Table 1). Of note, the mean age (standard deviation (*SD*)) at blood collection in the validation set was 55.5 (8.1) years and 53.1 (8.6) years in the discovery set. Consequently, the proportion of postmenopausal women was 61.8% in the validation set and 51.4% in the discovery set. In the validation set, 42.0% of participants had ever used oral contraceptive (vs 50.3% in the discovery set), 53.3% of women had received none or primary education (vs 47.3% in discovery set), 29.9% were physically inactive (vs 24.7% in discovery set), 16.9% were current smokers (vs 21.6% in discovery set), and 26.3% were alcohol non-consumers (vs 19.2% in discovery set).

In all participants ($N$ = 2358), strong correlations were observed between acyl-alkyl PCs (Fig. 1, Pearson's correlation coefficients 0.61 to 0.92), while moderate correlations were observed between acyl-alkyl PCs and PC aa C36:3 (0.41 to 0.55). Arginine was moderately correlated with all metabolites except for acetylcarnitine (C2), with an observed correlation of 0.19 with asparagine and correlations ranging from 0.11 to 0.13 with PCs. Asparagine showed similar low correlations (0.12 to 0.15) with PCs and a negative correlation with C2 (−0.17). C2 showed the greatest correlation with age (0.23), followed by PC aa C36:3 (0.19), while for other metabolites correlations with age ranged from −0.09 to 0.07.

In the discovery set, 104 associations (31% of the 336 associations tested, 8 metabolites × 42 variables) had *P*-values ≤0.05 (Supplementary Table 1, see Additional file 1). After correction of *P*-values for multiple testing, 57 of these associations remained significant (Table 2), which did not include any associations with arginine. Thirty associations were replicated in the validation set (same direction as in the discovery set, Supplementary Table 1, see Additional file 1) after Bonferroni correction of *P*-values, which did not include any associations with PC aa C36:3 (Table 2).

Figure 2 represents means of the metabolite concentrations across categories of variables in the overall population ($n$ = 2358), for metabolites and variables for which a significant association was detected in both the discovery and validation sets. Asparagine concentration was negatively associated with BMI, waist and hip circumferences, and WHR. C2 was positively associated with age but not with the other factors. PCs ae C36:2 and ae C38:2 were negatively associated with BMI, waist and hip circumferences, and waist/hip ratio. Negative associations with BMI, waist circumference, and waist/hip ratio were also observed for PCs ae C34:2 and ae C36:3. PC ae C34:2, C36:2, and 36:3 were additionally positively associated with total fat intake, and with saturated fatty acid intake, which was also positively associated with PC ae C38:2. For PC ae C36:2, additional associations were observed with alcohol intake at recruitment and over lifetime (negative) and with HLI and WCRF/AICR score (positive).

Analyses of interactions with BMI, menopausal, and fasting status (Supplementary Table 2, see Additional file 1) did not suggest any significant interaction with these variables in the associations reported above. The only interactions with significant *P*-values after correction for multiple testing were with menopausal status for the association between asparagine and age (P-int = 0.04) and with fasting status for the association of height and PC ae C38:2 (P-int = 0.03).

When excluding serum samples (restricting the analysis to plasma samples) from both discovery ($n$ = 40) and validation ($n$ = 88) sets, results were largely consistent with those in the main analyses (data not shown), except for generally larger *P*-values (due to the lower statistical power) that led to the following non-significant associations in the discovery set: asparagine and WCRF/AICR score, PC aa C36:3 and age and BMI, and PC ae C38:2 and trans-polyenoic fatty acid intake. In the validation set, associations between asparagine and hip and waist circumferences were not statistically significant anymore. However, estimates were very close in direction and magnitude to the ones obtained overall (before exclusion of serum samples).

After exclusion of participants with self-reported diabetes at blood collection (discovery set, $n$ = 45; validation set, $n$ = 26) or with missing information on diabetes (discovery set, $n$ = 86; validation set, $n$ = 74), associations were very similar in direction and magnitude

**Table 1** Main characteristics of women included (hormone non-users only), overall and in discovery and validation sets

| | Overall (*n* = 2358) | Discovery (*n* = 1572) | Validation (*n* = 786) |
|---|---|---|---|
| Age at blood collection (years) | 53.9 (8.5) | 53.1 (8.6) | 55.5 (8.1) |
| Fasting status at blood collection[a] (%) | | | |
| No | 936 (39.7) | 639 (40.6) | 297 (37.8) |
| In between | 375 (15.9) | 252 (16.0) | 123 (15.6) |
| Yes | 1047 (44.4) | 681 (43.3) | 366 (46.6) |
| Menopausal status at blood collection (%) | | | |
| Premenopausal | 722 (30.6) | 522 (33.2) | 200 (25.4) |
| Postmenopausal | 1294 (54.9) | 808 (51.4) | 486 (61.8) |
| Perimenopausal | 342 (14.5) | 242 (15.4) | 100 (12.7) |
| Age at first menstrual periods (years) (mean (*SD*)) | 13.1 (1.6) | 13.1 (1.6) | 13.1 (1.6) |
| Number of full-term pregnancies (mean (*SD*)) | 2.1 (1.3) | 2.0 (1.2) | 2.1 (1.4) |
| Age at first full-term pregnancy (years) (mean (*SD*)) | 25.2 (4.3) | 25.1 (4.4) | 25.4 (4.2) |
| Breastfeeding (in parous women) (%) | | | |
| Yes | 1669 (80.9) | 1110 (80.7) | 559 (81.2) |
| No | 280 (13.6) | 181 (13.2) | 99 (14.4) |
| Missing | 115 (5.6) | 85 (6.2) | 30 (4.4) |
| Ever used oral contraceptive (%) | 1120 (47.5) | 790 (50.3) | 330 (42.0) |
| Ever used MHT (%) | 297 (12.6) | 198 (12.6) | 99 (12.6) |
| Education level (%) | | | |
| Primary/no schooling | 1162 (49.3) | 743 (47.3) | 419 (53.3) |
| Technical/professional/secondary | 819 (34.7) | 560 (35.6) | 259 (33.0) |
| Longer education | 377 (16.0) | 269 (17.1) | 108 (13.7) |
| Physical activity (Cambridge Index) (%) | | | |
| Inactive | 623 (26.4) | 388 (24.7) | 235 (29.9) |
| Moderately inactive | 929 (39.4) | 623 (39.6) | 306 (38.9) |
| Moderately active | 450 (19.1) | 307 (19.5) | 143 (18.2) |
| Active | 356 (15.1) | 254 (16.2) | 102 (13.0) |
| Smoking status (%) | | | |
| Never | 1406 (59.6) | 934 (59.4) | 472 (60.1) |
| Former | 480 (20.4) | 299 (19.0) | 181 (23.0) |
| Smoker | 472 (20.0) | 339 (21.6) | 133 (16.9) |
| Alcohol consumption at recruitment (%) | | | |
| Non-drinker | 509 (21.6) | 302 (19.2) | 207 (26.3) |
| > 0–3 g/day | 707 (30.0) | 482 (30.7) | 225 (28.6) |
| > 3–12 g/day | 619 (26.3) | 416 (26.5) | 203 (25.8) |
| > 12–24 g/day | 337 (14.3) | 239 (15.2) | 98 (12.5) |
| > 24 g/day | 186 (7.9) | 133 (8.5) | 53 (6.7) |
| Height (cm) (mean (*SD*)) | 160.4 (6.8) | 160.6 (6.7) | 160.0 (6.8) |
| BMI (kg/m$^2$) (mean (*SD*)) | 26.0 (4.3) | 25.9 (4.3) | 26.3 (4.4) |
| Waist circumference (cm) (mean (*SD*)) | 81.9 (10.6) | 81.5 (10.4) | 82.6 (10.9) |
| Waist/hip ratio (mean (*SD*)) | 0.80 (0.07) | 0.80 (0.07) | 0.81 (0.07) |
| Total energy intake (kcal/day) (mean (*SD*)) | 2010.3 (547.7) | 2016.3 (557.6) | 1998.3 (527.4) |
| Healthy Lifestyle Index[b] (mean(*SD*)) | 12.6 (3.0) | 12.6 (3.0) | 12.8 (2.9) |
| WCRF/AICR score[c] (mean (*SD*)) | 3.9 (1.0) | 3.9 (1.0) | 3.9 (1.0) |

**Table 1** Main characteristics of women included (hormone non-users only), overall and in discovery and validation sets *(Continued)*

| | Overall (*n* = 2358) | Discovery (*n* = 1572) | Validation (*n* = 786) |
|---|---|---|---|
| Modified Mediterranean diet score (mean (*SD*)) | 4.3 (1.8) | 4.3 (1.8) | 4.5 (1.7) |
| Inflammatory score of the diet (mean (*SD*)) | 0.9 (1.7) | 0.89 (1.7) | 0.8 (1.7) |
| Diet Quality Index-International (mean (*SD*)) | 57.4 (7.8) | 57.2 (7.9) | 57.8 (7.7) |
| **Metabolite concentrations (normalized), μmol/L** | | | |
| Arginine (geometric mean (*SD*)) | 64.1 (2.6) | 64.2 (2.6) | 63.9 (2.6) |
| Asparagine (geometric mean (*SD*)) | 41.5 (2.7) | 41.4 (2.7) | 41.6 (2.7) |
| C2 (geometric mean (*SD*)) | 4.9 (2.7) | 4.9 (2.7) | 4.9 (2.6) |
| PC aa C36:3 (geometric mean (*SD*)) | 130.0 (2.6) | 127.0 (2.6) | 136.0 (2.7) |
| PC ae C34:2 (geometric mean (*SD*)) | 12.2 (2.7) | 12.1 (2.8) | 12.4 (2.7) |
| PC ae C36:2 (geometric mean (*SD*)) | 15.6 (2.7) | 15.5 (2.7) | 15.6 (2.7) |
| PC ae C36:3 (geometric mean (*SD*)) | 8.1 (2.7) | 8.0 (2.7) | 8.2 (2.7) |
| PC ae C38:2 (geometric mean (*SD*)) | 2.1 (2.7) | 2.1 (2.7) | 2.1 (2.8) |

*Abbreviations*: *AICR* American Institute for Cancer Research, *BMI* body mass index, *C2* acetylcarnitine, *MHT* menopause hormone therapy, *PC aa* phosphatidylcholine diacyl, *PC ae* phosphatidylcholine acyl-alkyl, *SD* standard deviation, *WCRF* World Cancer Research Fund
[a]No: < 3 h since last meal (< 4 h in Umeå); in between: 3–6 h since last meal (4–8 h in Umeå); yes: > 6 h since last meal (> 8 h in Umeå)
[b]Healthy Lifestyle Index was missing for 144 (6.1%) participants
[c]WCRF/AICR score was missing for 196 (8.3%) participants

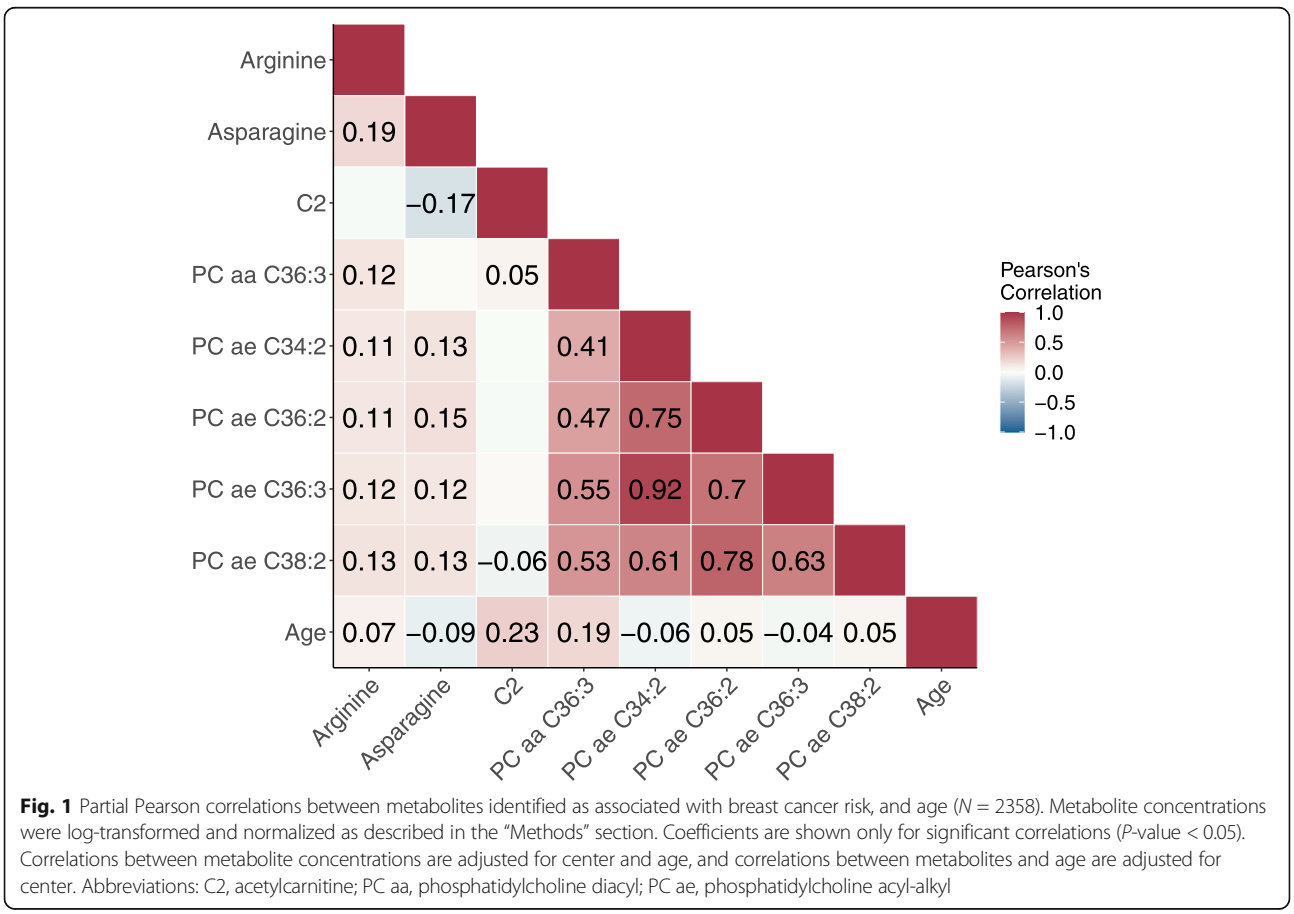



**Fig. 1** Partial Pearson correlations between metabolites identified as associated with breast cancer risk, and age (*N* = 2358). Metabolite concentrations were log-transformed and normalized as described in the "Methods" section. Coefficients are shown only for significant correlations (*P*-value < 0.05). Correlations between metabolite concentrations are adjusted for center and age, and correlations between metabolites and age are adjusted for center. Abbreviations: C2, acetylcarnitine; PC aa, phosphatidylcholine diacyl; PC ae, phosphatidylcholine acyl-alkyl

**Table 2** *P*-values for associations between metabolites and selected[a] variables

| Metabolite | Variable | Discovery | | Validation | |
|---|---|---|---|---|---|
| | | *P*-value | minP *P*-value | *P*-value | Bonf. *P*-value |
| Arginine[b] | - | - | - | - | - |
| Asparagine | **BMI** | **1.4E−19** | **1.4E−19** | **9.3E−07** | **5.6E−06** |
| | **Waist circumference** | **4.8E−16** | **4.8E−16** | **9.9E−05** | **6.0E−04** |
| | **Hip circumference** | **2.2E−13** | **2.2E−13** | **3.7E−03** | **2.2E−02** |
| | **Waist/hip ratio** | **1.3E−05** | **1.0E−03** | **8.2E−03** | **4.9E−02** |
| | Fiber intake | 1.3E−03 | 3.9E−02 | 1.5E−01 | 8.9E−01 |
| | WCRF/AICR score | 8.3E−04 | 2.4E−02 | 5.0E−01 | 1.0E+00 |
| C2 | **Age at blood collection** | **8.6E−04** | **2.6E−02** | **2.3E−04** | **1.1E−03** |
| | Alcohol consumption at recruitment | 2.8E−04 | 8.0E−03 | 2.0E−01 | 1.0E+00 |
| | Alcohol consumption (lifetime) | 8.6E−04 | 2.6E−02 | 3.5E−01 | 1.0E+00 |
| | Carbohydrate intake | 3.8E−05 | 1.0E−03 | 3.0E−01 | 1.0E+00 |
| | Total sugar intake | 2.1E−04 | 6.0E−03 | 1.1E−01 | 5.5E−01 |
| PC aa C36:3 | Age at blood collection | 1.0E−03 | 4.3E−02 | 1.2E−01 | 2.4E−01 |
| | BMI | 8.9E−04 | 3.6E−02 | 6.7E−02 | 1.3E−01 |
| PC ae C34:2 | Smoking status | 1.2E−04 | 1.0E−03 | 7.5E−01 | 1.0E+00 |
| | Smoking status and intensity | 2.0E−04 | 3.0E−03 | 8.4E−01 | 1.0E+00 |
| | **BMI** | **5.7E−10** | **45.7E−10** | **6.6E−05** | **7.9E−04** |
| | **Waist circumference** | **2.6E−14** | **2.6E−14** | **3.0E−07** | **3.6E−06** |
| | Hip circumference | 4.3E−05 | 4.3E−05 | 2.2E−02 | 2.6E−01 |
| | **Waist/hip ratio** | **5.0E−12** | **5.0E−12** | **3.3E−07** | **3.9E−06** |
| | Carbohydrate intake | 4.3E−06 | 4.3E−06 | 4.2E−02 | 5.0E−01 |
| | Total sugar intake | 1.4E−03 | 3.2E−02 | 9.8E−01 | 1.0E+00 |
| | **Total fat intake** | **2.2E−09** | **2.2E−09** | **1.4E−04** | **1.7E−03** |
| | **Fatty acids, total saturated intake** | **6.8E−10** | **6.8E−10** | **8.7E−07** | **1.0E−05** |
| | Fatty acids, total monounsaturated intake | 1.5E−04 | 1.0E−03 | 1.2E−01 | 1.0E+00 |
| | Modified Mediterranean diet score | 5.1E−04 | 1.5E−02 | 9.5E−03 | 1.1E−01 |
| PC ae C36:2 | Smoking status | 2.1E−04 | 6.0E−03 | 6.1E−01 | 1.0E+00 |
| | Smoking status and intensity | 2.4E−04 | 8.0E−03 | 9.0E−01 | 1.0E+00 |
| | **Alcohol consumption at recruitment** | **4.4E−08** | **4.4E−08** | **1.4E−04** | **1.7E−03** |
| | **Alcohol consumption (lifetime)** | **1.3E−06** | **1.3E−06** | **6.8E−04** | **8.1E−03** |
| | **BMI** | **4.4E−13** | **4.4E−13** | **7.0E−09** | **8.4E−08** |
| | **Waist circumference** | **2.1E−19** | **2.1E−19** | **1.8E−12** | **2.2E−11** |
| | **Hip circumference** | **1.1E−07** | **1.1E−07** | **7.7E−05** | **9.3E−04** |
| | **Waist/hip ratio** | **9.1E−15** | **9.1E−15** | **7.0E−10** | **8.4E−09** |
| | **Total fat intake** | **1.4E−11** | **1.4E−11** | **2.1E−05** | **2.5E−04** |
| | **Fatty acids, total saturated intake** | **2.2E−17** | **2.2E−17** | **1.6E−10** | **1.9E−09** |
| | **Healthy Lifestyle Index** | **2.8E−05** | **2.0E−03** | **2.3E−04** | **2.8E−03** |
| | **WCRF/AICR score** | **3.9E−05** | **2.0E−03** | **8.1E−05** | **9.7E−04** |
| PC ae C36:3 | Smoking status | 3.8E−04 | 7.0E−03 | 8.1E−01 | 1.0E+00 |
| | Smoking status and intensity | 1.4E−03 | 2.8E−02 | 8.6E−01 | 1.0E+00 |
| | **BMI** | **1.9E−07** | **1.9E−07** | **2.2E−03** | **2.7E−02** |
| | **Waist circumference** | **7.2E−12** | **7.2E−12** | **7.3E−05** | **8.8E−04** |
| | Hip circumference | 4.7E−04 | 8.0E−03 | 6.8E−02 | 8.1E−01 |

**Table 2** *P*-values for associations between metabolites and selected[a] variables *(Continued)*

| Metabolite | Variable | Discovery | | Validation | |
|---|---|---|---|---|---|
| | | *P*-value | minP *P*-value | *P*-value | Bonf. *P*-value |
| | **Waist/hip ratio** | **1.5E−10** | **1.5E−10** | **8.2E−05** | **9.8E−04** |
| | Carbohydrate intake | 1.9E−04 | 5.0E−03 | 2.2E−01 | 1.0E+00 |
| | **Total fat intake** | **8.5E−08** | **8.5E−08** | **4.1E−03** | **4.9E−02** |
| | **Fatty acids, total saturated intake** | **1.5E−06** | **1.5E−06** | **1.6E−03** | **2.0E−02** |
| | Fatty acids, total monounsaturated intake | 2.8E−05 | 2.8E−05 | 5.0E−02 | 6.0E−01 |
| | Fatty acids, total trans intake | 1.4E−04 | 3.0E−03 | 2.1E−01 | 1.0E+00 |
| | Modified Mediterranean diet score | 3.0E−04 | 6.0E−03 | 5.6E−02 | 6.7E−01 |
| PC ae C38:2 | Alcohol consumption at recruitment | 3.4E−04 | 2.4E−02 | 1.2E−02 | 9.9E−02 |
| | **BMI** | **6.2E−11** | **6.2E−11** | **1.9E−09** | **1.5E−08** |
| | **Waist circumference** | **1.6E−13** | **1.6E−13** | **2.5E−11** | **2.0E−10** |
| | **Hip circumference** | **7.6E−08** | **7.6E−08** | **1.2E−06** | **9.6E−06** |
| | **Waist/hip ratio** | **1.1E−07** | **1.1E−07** | **3.6E−06** | **2.9E−05** |
| | Total fat intake | 2.5E−04 | 1.7E−02 | 9.2E−03 | 7.3E−02 |
| | **Fatty acids, total saturated intake** | **1.7E−06** | **1.7E−06** | **1.2E−04** | **1.0E−03** |
| | Fatty acids, total trans-polyenoic intake | 1.2E−03 | 4.9E−02 | 6.3E−01 | 1.0E+00 |

Bold lines correspond to variable showing significant association after adjustment of *P*-value for multiple tests in both discovery and validation sets. Models were adjusted for center of recruitment, age, menopausal status (premenopausal, perimenopausal, postmenopausal), phase of the menstrual cycle for premenopausal women (follicular, ovulatory, luteal, missing), time of the day, and fasting status at blood collection (no, in between, yes)

*Abbreviations*: *AICR* American Institute for Cancer Research, *BMI* body mass index, *C2* acetylcarnitine, *PC aa* phosphatidylcholine diacyl, *PC ae* phosphatidylcholine acyl-alkyl, *WCRF* World Cancer Research Fund

[a]Only associations for which a significant *P*-value was detected after correction of *P*-values for multiple tests in the discovery sets are included

[b]No association was detected in the discovery set

to those observed in the whole dataset, although sometimes not significant in the validation set (data not shown), such as asparagine and hip circumference and waist/hip ratio, and PC ae C36:3 and BMI and total and saturated fat intakes.

## Discussion

In this study, we identified several lifestyle and anthropometric correlates of blood metabolites which have been previously associated with breast cancer risk in women not taking exogenous hormones at blood collection. Concentrations of PCs ae C34:2, ae C36:2, ae C36:3, and ae C38:2 showed negative associations with adiposity and positive associations with total (except for PC ae C38:2) and saturated fat intakes. PC ae C36:2 also showed a negative association with alcohol consumption and positive associations with the WCRF/AICR score and the Healthy Lifestyle Index. Asparagine concentrations were negatively associated with adiposity, and arginine concentrations were not associated with any of the variables examined. Acetylcarnitine concentrations were positively associated with age but not with any of the other factors. We did not identify any correlate of the only diacyl PC (PC aa C36:3) associated with breast cancer risk. These associations were consistent across different BMI, fasting status, and menopausal status categories.

Acyl-alkyl phosphatidylcholines have been previously associated with various lifestyle and dietary factors. In our work, concentrations of acyl-alkyl PCs were negatively associated with measures of adiposity (including BMI and waist circumference). This observation is consistent with the global pattern of negative associations between PCs ae and BMI previously reported in EPIC [12], in particular for PCs ae C38:2 [29] and ae C36:2 [30], and in the EPIC-Potsdam sub-cohort [31]. PC ae C38:2 and C34:2 were also associated with weight loss in an intervention study ($n$ = 17 participants) [12]. In the EPIC-Potsdam sub-cohort [32], a negative association of several PCs ae was reported with risk of type 2 diabetes, as well as a positive correlation with circulating high-density lipoprotein cholesterol. In an analysis of two studies of Japanese and American men and women [33], PCs ae C34:2, 36:3, and 38:2 were negatively associated with metabolic syndrome (in particular with high-density lipoprotein cholesterol and triglycerides), but not with elevated waist circumference. Among 200 Canadian adults younger than 55 years, concentrations of PCs ae C34:2, C36:2, and C36:3 were lower in obese participants with metabolic syndrome than in obese participants without metabolic syndrome and in normal weight participants [34], while an opposite trend was reported for several PCs aa. These results support an

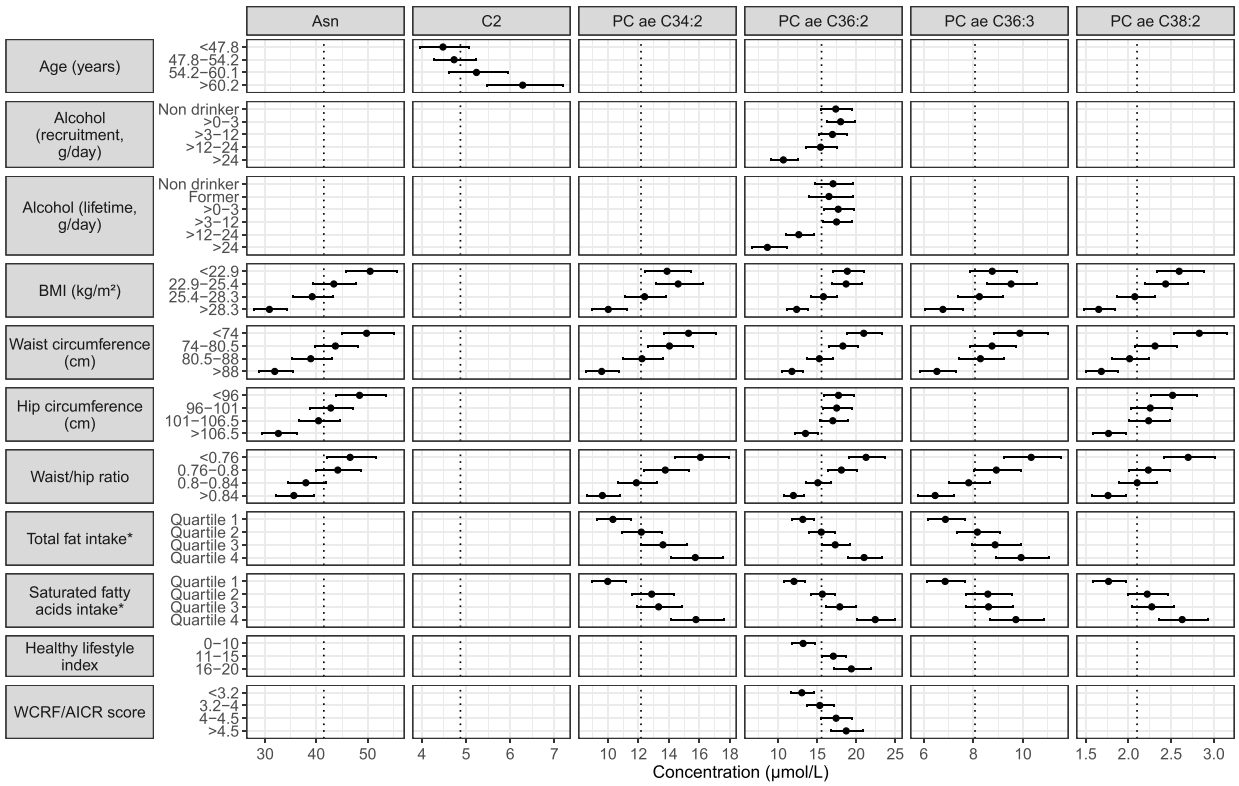



**Fig. 2** Adjusted means of metabolite concentrations by categories of correlates ($N$ = 2358). Only metabolites and variables for which a significant association was detected in the discovery and validation sets are shown. Adjusted means and their 95% confidence intervals were obtained from linear regression models adjusted for fasting status, center, age, date and time at blood collection, menopausal status, and phase of menstrual cycle at blood collection. Dotted lines indicate the overall means of metabolite concentration. *Residuals on total energy intake. Abbreviations: Asn, asparagine; AICR, American Institute for Cancer Research; BMI, body mass index; C2, acetylcarnitine; PC aa, phosphatidylcholine diacyl; PC ae, phosphatidylcholine acyl-alkyl; WCRF, World Cancer Research Fund

association of PCs with obesity or metabolic health that deserves further investigation.

Lower concentrations of PCs were reported in vegetarian and vegan men than in meat eaters [35]. Moreover, analyses in colorectal cancer patients (60% males) indicated positive associations of several PCs, mostly acyl-alkyl, with Western and carnivore dietary patterns [36]. These results are consistent with the positive association we report with saturated fat intake. However, few studies have been conducted in women, and an analysis conducted among healthy participants from the KarMeN study, not using exogenous hormones, suggested differences in plasma concentrations of some PCs between men and women, although PCs were not the most important components for predicting sex [37]. A recent metabolomic study of plasma lipid-related profiles and diet quality in the Nurses' Health Study [38] reported that PC C36:2 plasmalogen was associated with unhealthy components of the Alternate Healthy Eating Index.

A negative association of PCs and alcohol consumption, in particular PC ae C36:2, has been reported in EPIC, in both men and women [39]. A negative association with PC ae C36:2 was also observed separately in men and women from the KORA F4 study when comparing moderate-to-heavy drinkers (≥20 g/day for women, 40 g/day in men) with light drinkers (< 20 g/day for women, 40 g/day in men) [40], and in the CARLA study (men and women combined) [41].

The positive associations reported between PC ae C36:2 and the WCRF/AICR and HLI scores, which integrate alcohol and body weight components, likely reflect inverse associations of this metabolite with alcohol consumption and adiposity as demonstrated in the analyses of single correlates. These associations are in line with a recent study conducted in EPIC on metabolic signatures of a healthy lifestyle, assessed by the WCRF/AICR score [42]. In this work, PCs ae 36:2 and C38:2 were among endogenous metabolites with the greatest loadings (> 100 examined) in the signature of the WCRF/AICR score. This metabolic signature showed the greatest correlations with the recommendations regarding normal weight maintenance and alcohol avoidance, in line with the associations we report. In contrast, a study in

colorectal cancer patients indicated negative associations between several PCs ae and aa and the WCRF/AICR score [36]. However, the score was restricted to its dietary components, therefore not considering the body weight component.

Metabolomics studies on aging reported increasing circulating concentrations of acylcarnitines, mostly long-chain, with age [43, 44], which could reflect loss in mitochondrial function [45]. In a study [46] comparing metabolites in serum samples obtained 7 years apart from the same individuals (KORA S4 and KORA F4), acetylcarnitine and several other acylcarnitines increased in the follow-up samples compared with baseline samples. Associations of similar direction were observed in their validation study on samples collected 4 years apart, although not statistically significant after accounting for multiple testing. Acylcarnitines have also been associated with impaired glucose metabolism and insulin resistance, but these associations were most often reported for long-chain or odd short-chain acylcarnitines [47–51], although associations with acetylcarnitine (which is an even short-chain acylcarnitine) have also been reported [52]. In our previous work, this metabolite was the only one to show a positive association with breast cancer risk in age-matched cases and controls, suggesting that its association with age does not fully explain the association with breast cancer. In the present work, we did not observe any association of acetylcarnitine with anthropometric factors likely associated with metabolic health, in contrast with a positive association with BMI reported in the EPIC Norfolk cohort [53].

A negative association between circulating asparagine and obesity has been recently reported in different populations, including Europeans [50, 53], obese Iranian adults [54], and Japanese [55]. Negative associations with diabetes and coronary artery disease have also been reported [50, 53], in lean as well as in obese subjects [49]. However, most studies exploring the associations between amino acids and obesity showed significant associations only with branched-chain amino acids (which do not include asparagine) [49, 56]. Asparagine was also part of the metabolic signature of a healthy lifestyle derived in EPIC [42] and of the metabolic signature of BMI, waist circumference, and waist/hip ratio [12].

In our study, arginine was not associated with any of the factors investigated. This result contrasts with those in several studies reporting negative associations of arginine with age [46] and with obesity and alcohol intake, as well as a positive association with smoking in the EPIC Norfolk cohort [53], which however had not excluded hormone users. Arginine has also been negatively associated with hemoglobin concentrations and with insulin-like growth factor 1 and estradiol [57] in premenopausal women not using exogenous hormones. These observations may suggest that arginine concentrations could potentially be more tightly regulated by endogenous metabolism compared to lifestyle exposures.

Major strengths of this work include the wide variety of data collected which enabled us to investigate many potential correlates for the metabolites associated with breast cancer risk, and the large sample size of our study, compared to other metabolomics studies, where large studies are essential [58]. With the detailed information available on characteristics of women at blood collection, we were also able to exclude hormone users from our analysis, which is important as hormone use could possibly affect concentrations of some metabolites [59].

A first limitation to this work is the cross-sectional design, which prevents us from drawing any conclusions on the timing or causality of the associations. Another limitation is that the large sample size was achieved by pooling data from different previous studies, rather than by initial design, therefore adding methodological complexity because of analyses performed by different laboratories, with different instruments, and on different biological matrices. However, the analytical protocol used has shown high inter-laboratory reproducibility [60], and we addressed potential heterogeneity in metabolite concentrations by developing a dedicated pipeline [24] applied to the data prior to statistical analyses. In addition, for all metabolites included (except asparagine, not evaluated), high correlations were reported between measures in serum and in plasma ($r \geq 0.78$, except for arginine, $r = 0.50$), although concentrations were generally higher in serum than in plasma, in particular for arginine [61]. Good reliability of measurements was also reported for both matrices (intra-class correlations for the metabolites of interest ≥0.58 in plasma, ≥0.67 in serum) [62]. Furthermore, exclusion of serum samples did not substantially modify the results. A third limitation is the heterogeneity of fasting status of participants. However, variables to determine fasting status were carefully recorded, therefore enabling us to test the effect of this variable on the results, and we found no evidence of heterogeneity in the associations by fasting status. Dietary intakes were assessed using food frequency questionnaires adapted to local habits. These questionnaires were validated through a calibration approach using a common 24-h diet recall [63] to adjust for possible systematic misclassification in dietary measurements, and a validation study using 24-h urine samples was conducted [64]. Despite these methodological efforts, however, potential measurement error may persist because of recall bias, misreporting of consumption for certain foods, or errors related to the food composition tables used (despite careful matching [15]). Nevertheless, several cross-sectional studies showing good

correlations [65, 66] between intakes measured by food questionnaires and expected specific biomarkers suggest that data from food frequency questionnaires can be used for the purposes of the present work. Finally, the applied technology for PC measurement does not allow for precise identification of the compounds measured, since the signal observed is not specific and may correspond to different structural isomers. Further work is needed to investigate specifically associations with lipid compounds.

## Conclusions

In conclusion, this cross-sectional analysis identified several modifiable correlates of blood concentrations of metabolites associated with breast cancer risk. These associations may indicate possible mechanisms underlying associations between lifestyle and anthropometric factors, and risk of breast cancer. To better understand how our results could improve our current knowledge on the association between lifestyle factors and breast cancer risk, dedicated tools, such as mediation analysis, bring promising perspectives. Intervention studies would be required to evaluate the possible causality of the associations observed with modifiable factors and to assess whether concentrations of these specific metabolites could be modified through lifestyle changes.

### Abbreviations

AICR: American Institute for Cancer Research; Asn: Asparagine; BMI: Body mass index; C2: Acetylcarnitine; *CI*: Confidence interval; EPIC: European Prospective Investigation into Cancer and Nutrition; HLI: Healthy Lifestyle Index; IARC: International Agency for Research on Cancer; LLOQ: Lower limit of quantification; PC aa: Phosphatidylcholine diacyl; PC ae: Phosphatidylcholine acyl-alkyl; SD: Standard deviation; WCRF: World Cancer Research Fund

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-021-02183-2.

**Additional file 1:** Associations between metabolites and all variables tested as correlates, in discovery set and, for significant associations, in validation set (Supplementary Table 1); Stratified analyses by BMI, menopausal status, and fasting status at blood collection, for metabolites and variables showing a P-int < 0.20 after correction for multiple testing, in discovery set (Supplementary Table 2). Abbreviations: AICR American Institute for Cancer Research; Asn asparagine; BMI body mass index; C2 acetylcarnitine; HLI Healthy lifestyle index; PC aa phosphatidylcholine diacyl; PC ae phosphatidylcholine acyl-alkyl; SD Standard deviation; SE: Standard error of estimate; WCRF World Cancer Research Fund. (XLS 177 kb)

### Acknowledgements

The authors would like to thank Mr Bertrand Hemon for his support in preparing the databases, Ms Audrey Gicquiau and Dr David Achaintre for the analyses of samples in several of the original studies, and all EPIC participants. The EPIC-Norfolk team thank all the participants who have been part of the project and the many members of the study teams at the University of Cambridge who have enabled this research.

### IARC disclaimer

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

### Authors' contributions

The authors' responsibilities were as follows: SR, MH, and VV conceived and designed the research; MH analyzed the data with assistance of VV and SR; SR, MH, VV, LD, JAS, RCT, and MJG were responsible for drafting the manuscript; KO, CK, AT, LL, JAR, GS, TJ, VK, MBS, GM, SS, SP, RT, AM, JMAB, EMM, KSO, THN, TMS, AA, MJS, PA, SMCY, EA, LV, AW, AKH, EW, and IH provided the original data, information on the respective populations, and advice on the study design, analysis, and interpretation of the findings; all authors provided critical interpretation of the results and review of the first draft; all authors read and approved the final manuscript.

### Funding

This work was funded by the French National Cancer Institute (grant number 2015-166). Mathilde His' work reported here was undertaken during the tenure of a postdoctoral fellowship awarded by the International Agency for Research on Cancer, financed by the Fondation ARC.
The coordination of EPIC is financially supported by International Agency for Research on Cancer (IARC) and also by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, which has additional infrastructure support provided by the NIHR Imperial Biomedical Research Centre (BRC).
The national cohorts are supported by Danish Cancer Society (Denmark); Ligue 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); Dutch Ministry of Public Health, Welfare and Sports (VWS), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF) (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); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); and Cancer Research UK (14136 to EPIC-Norfolk (DOI 10.22025/2019.10.105.00004); C8221/A29017 to EPIC-Oxford), Medical Research Council (1000143, MR/N003284/1, MC-UU_12015/1 and MC_UU_00006/1 to EPIC-Norfolk; MR/M012190/1 to EPIC-Oxford) (UK). The funders were not involved in designing the study; collecting, analyzing, or interpreting the data; or writing or submitting the manuscript for publication.

### Availability of data and materials

EPIC data are available for investigators who seek to answer important questions on health and disease in the context of research projects that are consistent with the legal and ethical standard practices of IARC/WHO and the EPIC Centres. The primary responsibility for accessing the data belongs to the EPIC centers that provided them. For information on how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php.

## Declarations

### Ethics approval and consent to participate

This study complies with the Declaration of Helsinki. The study has been approved by the IARC Ethics Committee (ref IEC 14-02). Ethical review boards of IARC and all local institutions where participants had been recruited gave approval for the study, and all participants gave written informed consent for data collection and storage, as well as individual follow-up.

### Consent for publication

Not applicable.

## Competing interests

Marc J Gunter is an Editorial Adviser for BMC Medicine, and Joseph A Rothwell is an Editorial Board Member for BMC Medicine.
The other authors declare that they have no competing interests.

## Author details

[1]International Agency for Research on Cancer (IARC/WHO), Nutrition and Metabolism Branch, 150 cours Albert Thomas, 69372, CEDEX 08 Lyon, France. [2]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK. [3]Department of Public Health, Aarhus University, Aarhus, Denmark. [4]Danish Cancer Society Research Center, Copenhagen, Denmark. [5]Department of Public Health, Section of Environmental Health, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [6]Université Paris-Saclay, UVSQ, Inserm, CESP U1018, "Exposome and Heredity" team, Gustave Roussy, Villejuif, France. [7]Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Florence, Italy. [8]Department of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [9]Department of Molecular Epidemiology, German Institute of Human Nutrition, Nuthetal, Germany. [10]Institute of Nutritional Science, University of Potsdam, Potsdam, Germany. [11]Institute for Cancer Research, Prevention and Clinical Network (ISPRO), Florence, Italy. [12]Epidemiology and Prevention Unit, Fondazione IRCC S Instituto Nazionale dei Tumori di Milano, Milan, Italy. [13]Dipartimento Di Medicina Clinica E Chirurgia, Federico Ii University, Naples, Italy. [14]Cancer Registry and Histopathology Department, Provincial Health Authority (ASP 7) Ragusa, Ragusa, Italy. [15]Department of Clinical and Biological Sciences, University of Turin, Turin, Italy. [16]Center for Nutrition, Prevention, and Health Services, National Institute for Public Health and the Environment (RIVM), Bilthoven 3720, BA, the Netherlands. [17]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands. [18]Department of Community Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway. [19]Unit of Nutrition and Cancer, Catalan Institute of Oncology - ICO, L'Hospitalet de Llobregat, Spain. [20]Nutrition and Cancer Group; Epidemiology, Public Health, Cancer Prevention and Palliative Care Program; Bellvitge Biomedical Research Institute - IDIBELL, L'Hospitalet de Llobregat, Spain. [21]Escuela Andaluza de Salud Pública (EASP), Granada, Spain. [22]Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain. [23]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. [24]Department of Preventive Medicine and Public Health, University of Granada, Granada, Spain. [25]Ministry of Health of the Basque Government, Sub-Directorate for Public Health and Addictions of Gipuzkoa, San Sebastián, Spain. [26]Biodonostia Health Research Institute, Group of Epidemiology of Chronic and Communicable Diseases, San Sebastián, Spain. [27]CIBER Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III (ISCIII), Madrid, Spain. [28]Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain. [29]Research Group on Demography and Health, National Faculty of Public Health, University of Antioquia, Medellín, Colombia. [30]Navarra Public Health Institute, Pamplona, Spain. [31]IdiSNA, Navarra Institute for Health Research, Pamplona, Spain. [32]Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden. [33]Sustainable Health, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden. [34]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK. [35]International Agency for Research on Cancer (IARC/WHO), Office of the Director, Lyon, France.

Received: 9 July 2021 Accepted: 9 November 2021
Published online: 10 December 2021

## References

1. 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(1):178. https://doi.org/10.1186/s12916-019-1408-4.
2. Lecuyer L, Dalle C, Lyan B, Demidem A, Rossary A, Vasson MP, et al. Plasma metabolomic signatures associated with long-term breast cancer risk in the SU.VI.MAX Prospective Cohort. Cancer Epidemiol Biomarkers Prev. 2019;28(8):1300–7. https://doi.org/10.1158/1055-9965.EPI-19-0154.
3. Lecuyer L, Victor Bala A, Deschasaux M, Bouchemal N, Nawfal Triba M, Vasson MP, et al. NMR metabolomic signatures reveal predictive plasma metabolites associated with long-term risk of developing breast cancer. Int J Epidemiol. 2018;47(2):484–94. https://doi.org/10.1093/ije/dyx271.
4. Moore SC, Mazzilli KM, Sampson JN, Matthews CE, Carter BD, Playdon MC, et al. A Metabolomics Analysis of Postmenopausal Breast Cancer Risk in the Cancer Prevention Study II. Metabolites. 2021;11(2). https://doi.org/10.3390/metabo11020095.
5. Moore SC, Playdon MC, Sampson JN, Hoover RN, Trabert B, Matthews CE, et al. A metabolomics analysis of body mass index and postmenopausal breast cancer risk. J Natl Cancer Inst. 2018;110(6):588–97. https://doi.org/10.1093/jnci/djx244.
6. Playdon MC, Ziegler RG, Sampson JN, Stolzenberg-Solomon R, Thompson HJ, Irwin ML, et al. Nutritional metabolomics and breast cancer risk in a prospective study. Am J Clin Nutr. 2017;106(2):637–49. https://doi.org/10.3945/ajcn.116.150912.
7. Jobard E, Dossus L, Baglietto L, Fornili M, Lecuyer L, Mancini FR, et al. Investigation of circulating metabolites associated with breast cancer risk by untargeted metabolomics: a case-control study nested within the French E3N cohort. Br J Cancer. 2021;124(10):1734–43. https://doi.org/10.1038/s41416-021-01304-1.
8. Kuhn T, Floegel A, Sookthai D, Johnson T, Rolle-Kampczyk U, Otto W, et al. Higher plasma levels of lysophosphatidylcholine 18:0 are related to a lower risk of common cancers in a prospective metabolomics study. BMC Med. 2016;14(1):13. https://doi.org/10.1186/s12916-016-0552-3.
9. Wishart DS. Metabolomics for investigating physiological and pathophysiological processes. Physiol Rev. 2019;99(4):1819–75. https://doi.org/10.1152/physrev.00035.2018.
10. 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(6B):1113–24. https://doi.org/10.1079/PHN2002394.
11. Dossus L, Kouloura E, Biessy C, Viallon V, Siskos AP, Dimou N, et al. Prospective analysis of circulating metabolites and endometrial cancer risk. Gynecol Oncol. 2021;162(2):475–81. https://doi.org/10.1016/j.ygyno.2021.06.001.
12. 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(1):101. https://doi.org/10.1186/s12916-021-01970-1.
13. Johansson M, Fanidi A, Muller DC, Bassett JK, Midttun O, Vollset SE, et al. Circulating Biomarkers of One-Carbon Metabolism in Relation to Renal Cell Carcinoma Incidence and Survival. J Natl Cancer Inst. 2014;106(12). https://doi.org/10.1093/jnci/dju327.
14. Stepien M, Duarte-Salles T, Fedirko V, Floegel A, Barupal DK, Rinaldi S, et al. Alteration of amino acid and biogenic amine metabolism in hepatobiliary cancers: findings from a prospective cohort study. Int J Cancer. 2016;138(2):348–60. https://doi.org/10.1002/ijc.29718.
15. Van Puyvelde H, Perez-Cornago A, Casagrande C, Nicolas G, Versele V, Skeie G, et al. Comparing Calculated Nutrient Intakes Using Different Food Composition Databases: Results from the European Prospective Investigation into Cancer and Nutrition (EPIC) Cohort. Nutrients. 2020;12(10). https://doi.org/10.3390/nu12102906.
16. Spencer EA, Appleby PN, Davey GK, Key TJ. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. Public Health Nutr. 2002;5(4):561–5. https://doi.org/10.1079/PHN2001322.
17. Tehard B, van Liere MJ, Com NC, Clavel-Chapelon F. Anthropometric measurements and body silhouette of women: validity and perception. J Am Diet Assoc. 2002;102(12):1779–84. https://doi.org/10.1016/S0002-8223(02)90381-0.
18. Agudo A, Cayssials V, Bonet C, Tjonneland A, Overvad K, Boutron-Ruault MC, et al. Inflammatory potential of the diet and risk of gastric cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Am J Clin Nutr. 2018;107(4):607–16. https://doi.org/10.1093/ajcn/nqy002.
19. Trichopoulou A, Orfanos P, Norat T, Bueno-de-Mesquita B, Ocke MC, Peeters PH, et al. Modified Mediterranean diet and survival: EPIC-elderly prospective cohort study. BMJ. 2005;330(7498):991. https://doi.org/10.1136/bmj.38415.644155.8F.
20. Lassale C, Gunter MJ, Romaguera D, Peelen LM, Van der Schouw YT, Beulens JW, et al. Diet quality scores and prediction of all-cause, cardiovascular and cancer mortality in a pan-European cohort study. PLoS One. 2016;11(7):e0159025. https://doi.org/10.1371/journal.pone.0159025.

21. Kim S, Haines PS, Siega-Riz AM, Popkin BM. The Diet Quality Index-International (DQI-I) provides an effective tool for cross-national comparison of diet quality as illustrated by China and the United States. J Nutr. 2003; 133(11):3476–84. https://doi.org/10.1093/jn/133.11.3476.
22. McKenzie F, Biessy C, Ferrari P, Freisling H, Rinaldi S, Chajes V, et al. Healthy lifestyle and risk of cancer in the European Prospective Investigation Into Cancer and Nutrition cohort study. Medicine (Baltimore). 2016;95(16):e2850. https://doi.org/10.1097/MD.0000000000002850.
23. Romaguera D, Vergnaud AC, Peeters PH, van Gils CH, Chan DS, Ferrari P, 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(1):150–63. https://doi.org/10.3945/ajcn.111.031674.
24. 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(9). https://doi.org/10.3390/metabo11090631.
25. Kaaks R, Berrino F, Key T, Rinaldi S, Dossus L, Biessy C, et al. Serum sex steroids in premenopausal women and breast cancer risk within the European Prospective Investigation into Cancer and Nutrition (EPIC). J Natl Cancer Inst. 2005;97(10):755–65. https://doi.org/10.1093/jnci/dji132.
26. 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(4):407–13. https://doi.org/10.1079/PHN2002439.
27. Westfall PH, Young SS. Resampling-based multiple testing: examples and methods for p-value adjustment. New York: John Wiley & Sons; 1993.
28. Loftfield E, Stepien M, Viallon V, Trijsburg L, Rothwell JA, Robinot N, et al. Novel biomarkers of habitual alcohol intake and associations with risk of pancreatic and liver cancers and liver disease mortality. J Natl Cancer Inst. 2021;113(11):1542–50. https://doi.org/10.1093/jnci/djab078.
29. 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(9):3137–46. https://doi.org/10.1021/acs.jproteome.6b01062.
30. Assi N, Thomas DC, Leitzmann M, Stepien M, Chajes 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(5):531–40. https://doi.org/10.1158/1055-9965.EPI-17-0649.
31. 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(11):1388–96. https://doi.org/10.1038/ijo.2014.39.
32. Floegel A, Stefan N, Yu Z, Muhlenbruch 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(2):639–48. https://doi.org/10.2337/db12-0495.
33. Mahajan UV, Varma VR, Huang CW, An Y, Tanaka T, Ferrucci L, et al. Blood Metabolite Signatures of Metabolic Syndrome in Two Cross-Cultural Older Adult Cohorts. Int J Mol Sci. 2020;21(4). https://doi.org/10.3390/ijms21041324.
34. Allam-Ndoul B, Guenard F, Garneau V, Cormier H, Barbier O, Perusse L, et al. Association between Metabolite Profiles, Metabolic Syndrome and Obesity Status. Nutrients. 2016;8(6). https://doi.org/10.3390/nu8060324.
35. Schmidt JA, Rinaldi S, Ferrari P, Carayol M, Achaintre D, Scalbert A, et al. Metabolic profiles of male meat eaters, fish eaters, vegetarians, and vegans from the EPIC-Oxford cohort. Am J Clin Nutr. 2015;102(6):1518–26. https://doi.org/10.3945/ajcn.115.111989.
36. Geijsen A, Kok DE, van Zutphen M, Keski-Rahkonen P, Achaintre D, Gicquiau A, et al. Diet quality indices and dietary patterns are associated with plasma metabolites in colorectal cancer patients. Eur J Nutr. 2021;60(6):3171–84. https://doi.org/10.1007/s00394-021-02488-1.
37. Rist MJ, Roth A, Frommherz L, Weinert CH, Kruger R, Merz B, et al. Metabolite patterns predicting sex and age in participants of the Karlsruhe Metabolomics and Nutrition (KarMeN) study. PLoS One. 2017;12(8): e0183228. https://doi.org/10.1371/journal.pone.0183228.
38. Bagheri M, Willett W, Townsend MK, Kraft P, Ivey KL, Rimm EB, et al. A lipid-related metabolomic pattern of diet quality. Am J Clin Nutr. 2020;112(6): 1613–30. https://doi.org/10.1093/ajcn/nqaa242.
39. van Roekel EH, Trijsburg L, Assi N, Carayol M, Achaintre D, Murphy N, et al. Circulating Metabolites Associated with Alcohol Intake in the European Prospective Investigation into Cancer and Nutrition Cohort. Nutrients. 2018; 10(5). https://doi.org/10.3390/nu10050654.
40. Jaremek M, Yu Z, Mangino M, Mittelstrass K, Prehn C, Singmann P, et al. Alcohol-induced metabolomic differences in humans. Transl Psychiatry. 2013;3(7):e276. https://doi.org/10.1038/tp.2013.55.
41. Lacruz ME, Kluttig A, Tiller D, Medenwald D, Giegling I, Rujescu D, et al. Cardiovascular risk factors associated with blood metabolite concentrations and their alterations during a 4-year period in a population-based cohort. Circ Cardiovasc Genet. 2016;9(6):487–94. https://doi.org/10.1161/CIRCGENETICS.116.001444.
42. Rothwell JA, Murphy N, Besevic 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. 2020. https://doi.org/10.1016/j.cgh.2020.11.045.
43. Jarrell ZR, Smith MR, Hu X, Orr M, Liu KH, Quyyumi AA, et al. Plasma acylcarnitine levels increase with healthy aging. Aging (Albany NY). 2020; 12(13):13555–70. https://doi.org/10.18632/aging.103462.
44. Srivastava S. Emerging Insights into the Metabolic Alterations in Aging Using Metabolomics. Metabolites. 2019;9(12). https://doi.org/10.3390/metabo9120301.
45. Flanagan JL, Simmons PA, Vehige J, Willcox MD, Garrett Q. Role of carnitine in disease. Nutr Metab (Lond). 2010;7(1):30. https://doi.org/10.1186/1743-7075-7-30.
46. Chak CM, Lacruz ME, Adam J, Brandmaier S, Covic M, Huang J, et al. Ageing Investigation Using Two-Time-Point Metabolomics Data from KORA and CARLA Studies. Metabolites. 2019;9(3). https://doi.org/10.3390/metabo9030044.
47. Pallares-Mendez R, Aguilar-Salinas CA, Cruz-Bautista I, Del Bosque-Plata L. Metabolomics in diabetes, a review. Ann Med. 2016;48(1-2):89–102. https://doi.org/10.3109/07853890.2015.1137630.
48. Mihalik SJ, Goodpaster BH, Kelley DE, Chace DH, Vockley J, Toledo FG, et al. Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipotoxicity. Obesity (Silver Spring). 2010; 18(9):1695–700. https://doi.org/10.1038/oby.2009.510.
49. Libert DM, Nowacki AS, Natowicz MR. Metabolomic analysis of obesity, metabolic syndrome, and type 2 diabetes: amino acid and acylcarnitine levels change along a spectrum of metabolic wellness. PeerJ. 2018;6:e5410. https://doi.org/10.7717/peerj.5410.
50. Ottosson F, Smith E, Melander O, Fernandez C. Altered asparagine and glutamate homeostasis precede coronary artery disease and type 2 diabetes. J Clin Endocrinol Metab. 2018;103(8):3060–9. https://doi.org/10.1210/jc.2018-00546.
51. Rangel-Huerta OD, Pastor-Villaescusa B, Gil A. Are we close to defining a metabolomic signature of human obesity? A systematic review of metabolomics studies. Metabolomics. 2019;15(6):93. https://doi.org/10.1007/s11306-019-1553-y.
52. Adams SH, Hoppel CL, Lok KH, Zhao L, Wong SW, Minkler PE, et al. Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid beta-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. J Nutr. 2009;139(6):1073–81. https://doi.org/10.3945/jn.108.103754.
53. Pietzner M, Stewart ID, Raffler J, Khaw KT, Michelotti GA, Kastenmuller G, et al. Plasma metabolites to profile pathways in noncommunicable disease multimorbidity. Nat Med. 2021;27(3):471–9. https://doi.org/10.1038/s41591-021-01266-0.
54. Bagheri M, Djazayery A, Farzadfar F, Qi L, Yekaninejad MS, Aslibekyan S, et al. Plasma metabolomic profiling of amino acids and polar lipids in Iranian obese adults. Lipids Health Dis. 2019;18(1):94. https://doi.org/10.1186/s12944-019-1037-0.
55. Takashina C, Tsujino I, Watanabe T, Sakaue S, Ikeda D, Yamada A, et al. Associations among the plasma amino acid profile, obesity, and glucose metabolism in Japanese adults with normal glucose tolerance. Nutr Metab (Lond). 2016;13(1):5. https://doi.org/10.1186/s12986-015-0059-5.
56. Bi X, Henry CJ. Plasma-free amino acid profiles are predictors of cancer and diabetes development. Nutr Diabetes. 2017;**7**(3):e249. https://doi.org/10.1038/nutd.2016.55.
57. Nagata C, Wada K, Tsuji M, Hayashi M, Takeda N, Yasuda K. Plasma amino acid profiles are associated with biomarkers of breast cancer risk in

premenopausal Japanese women. Cancer Causes Control. 2014;25(2):143–9. https://doi.org/10.1007/s10552-013-0316-8.
58. Moore SC. Metabolomics and breast cancer: scaling up for robust results. BMC Med. 2020;18(1):18. https://doi.org/10.1186/s12916-019-1484-5.
59. Ruoppolo M, Campesi I, Scolamiero E, Pecce R, Caterino M, Cherchi S, et al. Serum metabolomic profiles suggest influence of sex and oral contraceptive use. Am J Transl Res. 2014;6(5):614–24.
60. Siskos AP, Jain P, Romisch-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(1):656–65. https://doi.org/10.1021/acs.analchem.6b02930.
61. Yu Z, Kastenmuller G, He Y, Belcredi P, Moller G, Prehn C, et al. Differences between human plasma and serum metabolite profiles. PLoS One. 2011; 6(7):e21230. https://doi.org/10.1371/journal.pone.0021230.
62. Breier M, Wahl S, Prehn C, Fugmann M, Ferrari U, Weise M, et al. Targeted metabolomics identifies reliable and stable metabolites in human serum and plasma samples. PLoS One. 2014;9(2):e89728. https://doi.org/10.1371/journal.pone.0089728.
63. 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(6B):1125–45. https://doi.org/10.1079/PHN2002395.
64. Slimani N, Bingham S, Runswick S, Ferrari P, Day NE, Welch AA, et al. Group level validation of protein intakes estimated by 24-hour diet recall and dietary questionnaires against 24-hour urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study. Cancer Epidemiol Biomarkers Prev. 2003;12(8):784–95.
65. Saadatian-Elahi M, Slimani N, Chajes V, Jenab M, Goudable J, Biessy C, et al. Plasma phospholipid fatty acid profiles and their association with food intakes: results from a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition. Am J Clin Nutr. 2009;89(1):331–46. https://doi.org/10.3945/ajcn.2008.26834.
66. Al-Delaimy WK, Ferrari P, Slimani N, Pala V, Johansson I, Nilsson S, et al. Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). Eur J Clin Nutr. 2005;59(12):1387–96. https://doi.org/10.1038/sj.ejcn.1602252.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

**Learn more** biomedcentral.com/submissions

