



# Diet quality indices and dietary patterns are associated with plasma metabolites in colorectal cancer patients

Anne J. M. R. Geijsen<sup>1</sup> · Dieuwertje E. Kok<sup>1</sup> · Moniek van Zutphen<sup>1</sup> · Pekka Keski-Rahkonen<sup>2</sup> · David Achaintre<sup>2</sup> · Audrey Gicquiau<sup>2</sup> · Andrea Gsur<sup>3</sup> · Flip M. Kruijt<sup>4</sup> · Cornelia M. Ulrich<sup>5,6</sup> · Matty P. Weijenberg<sup>7</sup> · Johannes H. W. de Wilt<sup>8</sup> · Evertine Wesselink<sup>1</sup> · Augustin Scalbert<sup>2</sup> · Ellen Kampman<sup>1</sup> · Fränzel J. B. van Duijnhoven<sup>1</sup>

Received: 8 May 2020 / Accepted: 8 January 2021 / Published online: 5 February 2021  
© The Author(s) 2021

## Abstract

**Purpose** Emerging evidence suggests that diet is linked to survival in colorectal cancer patients, although underlying mechanisms are not fully understood. The aim of this study was to evaluate whether dietary exposures are associated with metabolite concentrations in colorectal cancer patients.

**Methods** Concentrations of 134 metabolites of the Biocrates Absolute<sup>IDQ</sup> p180 kit were quantified in plasma samples collected at diagnosis from 195 stage I-IV colorectal cancer patients. Food frequency questionnaires were used to calculate adherence to the World Cancer Research Fund (WCRF) dietary recommendations and the Dutch Healthy Diet (DHD15) index as well as to construct dietary patterns using Principal Component Analysis. Multivariable linear regression models were used to determine associations between dietary exposures and metabolite concentrations. All models were adjusted for age, sex, body mass index, smoking status, analytical batch, cancer stage, and multiple testing using false discovery rate.

**Results** Participants had a mean (SD) age of 66 (9) years, were mostly men (60%), and mostly diagnosed with stage II and III cancer. For the dietary pattern analyses, Western, Carnivore, and Prudent patterns were identified. Better adherence to the WCRF dietary recommendations was associated with lower concentrations of ten phosphatidylcholines. Higher intake of the Carnivore pattern was associated with higher concentrations of two phosphatidylcholines. The DHD15-index, Western pattern, or Prudent pattern were not associated with metabolite concentrations.

**Conclusion** In the current study, the WCRF dietary score and the Carnivore pattern are associated with phosphatidylcholines. Future research should elucidate the potential relevance of phosphatidylcholine metabolism in the colorectal cancer continuum.

**Clinical trial registry** ClinicalTrials.gov Identifier: NCT03191110.

**Keywords** Colorectal cancer patients · Diet quality indices · Dietary patterns · Metabolites · Metabolomics

✉ Fränzel J. B. van Duijnhoven  
franzel.vanduijnhoven@wur.nl

<sup>1</sup> Division of Human Nutrition and Health, Wageningen University and Research, P.O. Box 17, 6700 AA Wageningen, The Netherlands  
<sup>2</sup> Biomarker Group, International Agency for Research on Cancer, Lyon, France  
<sup>3</sup> Institute of Cancer Research, Department of Medicine I, Medical University of Vienna, Vienna, Austria  
<sup>4</sup> Department of Surgery, Hospital Gelderse Vallei, Ede, The Netherlands

<sup>5</sup> Huntsman Cancer Institute, Salt Lake City, Utah, USA

<sup>6</sup> Department of Population Health Sciences, University of Utah, Salt Lake City, Utah, USA

<sup>7</sup> Department of Epidemiology, GROW School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands

<sup>8</sup> Department of Surgery, Division of Surgical Oncology and Gastrointestinal Surgery, Radboud University Medical Centre, Nijmegen, The Netherlands

## Abbreviations

AICR	American Institute for Cancer Research
BMI	Body mass index
DHD15-index	Dutch Healthy Diet index 2015
FDR	False Discovery Rate
FFQ	Food Frequency Questionnaire
FIA-MS/MS	Flow injection analysis-tandem mass spectrometry
IARC	International Agency for Research on Cancer
LOD	Limit of detection
LysoPC a	Lysophosphatidylcholine (acyl)
SD	Standard deviation
SM (OH)	(Hydroxy) sphingomyelin
SQUASH	Short QUestionnaire to ASsess Health enhancing physical activity
PC aa	Phosphatidylcholine (diacyl)
PC ae	Phosphatidylcholine (acyl-alkyl)
UHPLC-MS/MS	Ultra-high performance liquid chromatography-tandem mass spectrometry
WCRF	World Cancer Research Fund

## Introduction

Previous research suggests that there is a relation between diet and survival after colorectal cancer diagnosis [1]. However, the underlying mechanisms are largely unknown. The identification of metabolites associated with diet in colorectal cancer patients could be a first step in unravelling the link between dietary exposures and colorectal cancer progression [2–4] and understanding the biological processes involved [5, 6]. Metabolomics measures a range of small molecules, many of which belong to a number of different biochemical pathways. Such metabolic profiles can provide a snapshot of the current metabolic state of the body, characteristic of a phenotype [7] and are, therefore, increasingly used to study the interface of diet, lifestyle and diseases [8–10].

Thus far, research has shown that dietary exposures can be associated with metabolite concentrations in blood. Pre-defined diet quality indices [11] as well as distinct dietary patterns such as veganism [12, 13] were reported to be associated with specific blood metabolites. Three diet quality indices were associated with metabolites, including mainly lipids and amino acids [11]. Another study reported that a vegan diet was associated with lower concentrations of glycerophospholipids, sphingolipids and amino acids compared to a diet containing meat and/or fish [12].

In terms of understanding the complex relationship between diet and metabolites, investigating dietary patterns and indices, instead of single nutrients or food groups, is of specific interest since nutrients and food groups may interact with each other [14]. Diet quality indices and dietary

patterns are used to assess exposure to combinations of food groups. An example of a diet quality index, commonly used to investigate the relationship with health outcomes after cancer diagnosis [15–17], is the World Cancer Research Fund (WCRF) score that is based on the 2018 cancer prevention recommendations of the WCRF/American Institute for Cancer Research (AICR) [18]. The Dutch Healthy Diet index (DHD15-index) is used to assess the adherence to the Dutch dietary guidelines of 2015 [19] and has also previously been associated with various health outcomes [20, 21]. In addition to these indices, dietary patterns are data-driven and are also commonly evaluated against health outcomes after cancer diagnosis [22].

It is well-established that disease can have a great impact on metabolism [23–25] and, to the best of our knowledge, no research has been conducted investigating the associations between dietary exposures and metabolites in cancer patients. Metabolites associated with dietary exposures in colorectal cancer patients may give clues to potential underlying mechanisms for colorectal cancer progression which could be studied in detail in the future. Therefore, the aim of this explorative study was to investigate whether the diet, evaluated using diet quality indices and dietary patterns, is associated with plasma metabolites in colorectal cancer patients.

## Methods

### Study population

In total, 200 stage I-IV colorectal cancer patients with available plasma metabolite concentrations of the COLON study [26], a prospective cohort study among colorectal cancer patients in the Netherlands, were considered for the present study. The design and recruitment of the COLON study has been described earlier [26]. Participants were recruited from 11 hospitals in the Netherlands, shortly after colorectal cancer diagnosis. Females and males of all ages and of any stage of the disease were eligible. Non-Dutch speaking patients, patients with a history of colorectal cancer or (partial) bowel resection, chronic inflammatory bowel disease, hereditary colorectal cancer syndromes (e.g. Lynch syndrome, Familial Adenomatous Polyposis, Peutz-Jegher), dementia or another mental condition causing an inability to fill out a questionnaire correctly were excluded. All participants provided a written informed consent. The COLON study was approved by the Committee on Research involving Human Subjects (region Arnhem-Nijmegen), the Netherlands.

Participants with missing dietary intake data ( $n = 2$ ) or with a missing cancer stage ( $n = 3$ ) were excluded from the current study, resulting in a final study population of  $n = 195$  stage I-IV colorectal cancer patients for analysis.

## Data collection

Habitual dietary intake in the month prior to diagnosis was assessed using a 204-item validated, semi-quantitative food frequency questionnaire (FFQ) developed by the Division of Human Nutrition and Health of Wageningen University & Research, the Netherlands [27, 28]. The FFQ was used to calculate a priori defined diet quality indices and to construct a posteriori data-driven dietary patterns. Demographic and lifestyle characteristics such as sex, age, weight, height, and smoking habits were assessed using self-administered questionnaires. All questionnaires were filled out prior to tumor resection. Medical information, including cancer stage, tumor location, and treatment strategies, was collected using the Dutch ColoRectal Audit [29].

Non-fasted plasma EDTA samples were collected upon recruitment, which were intended before the start of treatment, and stored at  $-80^{\circ}\text{C}$  using a standardized protocol to ensure identical sample handling across the eleven hospitals.

## Diet quality indices

Two diet quality indices have been included in the current study, namely the WCRF dietary score and the DHD15-index. Briefly, the WCRF dietary score is based on the 2018 WCRF/AICR recommendations for cancer prevention using the standard WCRF/AICR score developed by Shams-White et al. [18]. Since the current study focusses on dietary intake, the recommendations regarding weight, physical activity, supplement use, and breastfeeding were not included. The remaining recommendations were: (1) eat a diet rich in whole grains, vegetables, fruits, and beans, (2) limit consumption of ‘fast foods’ and other processed foods high in fat, starches or sugar, (3) limit consumption of red and processed meat, (4) limit consumption of sugar-sweetened drinks, and (5) reduce alcohol consumption. Quantitative criteria were used as cut-off points for all recommendations, except for the recommendation (2) limit consumption of ‘fast foods’ and other processed foods high in fat, starches or sugar, where cut-offs were based on tertiles calculated as a percentage of total energy intake from processed foods. Processed foods included French fries, crisps, pastry and biscuits, savory snacks, sugar and candy, sauces, pizza, pancake, sandwich fillings high in sugar or fat, refined grain products, and sweet dairy desserts. Processed meat included sausages, bacon, ribs, ham, cold cuts, and unknown types of meat. Sugary drinks included sugar-sweetened soft drinks, sugar-sweetened dairy drinks, and fruit juices. We did not include yoghurt and cheese, nuts, oils and fats, and diet soft drinks in the WCRF dietary score, since these food or food groups are not part of the WCRF recommendations.

The score assigned for each recommendation of the WCRF dietary score was 1 when the recommendation was

met (full adherence), a score of 0.5 was assigned to moderate adherence and a score of 0 was assigned to low adherence. The recommendation regarding a diet rich in whole grains, vegetables, fruit, and beans, included sub-recommendations for fiber intake and for fruit and vegetable consumption. As a result, the recommendation score was the sum of sub-recommendation scores of fiber intake and fruit and vegetables intake, meaning that plausible scores were 0, 0.25, 0.5, 0.75, and 1. The overall score of the WCRF dietary score ranged from 0 to 5.

The DHD15-index [19] was developed on the basis of the 2015 Dutch dietary guidelines [30] and refers to 15 recommendations. In the current study, the recommendations regarding coffee consumption and sodium intake were excluded since the type of coffee and sodium intake were not assessed in the COLON study. The DHD15-index used in the current study included the intake of sugary drinks, liquid fats and oils, processed meat, red meat, nuts, dairy products, refined grains products, whole grain products, vegetables, alcohol, legumes, solid fat, fruit, fish, and tea. The scores for each individual recommendation ranged from 0 to 10 points with a maximum total DHD15-index score of 130 points.

For both indices, a higher score represents a healthier diet, i.e. a better compliance with the recommendations of the corresponding diet quality index. Details on the used diet quality indices have been described before [18, 19, 31].

## Empirical construction of dietary patterns

Total intake of food items (g/d) and total energy intake (kcal/d) were calculated based on frequency of intake, number of portions, portion size, and the type of products, as recorded in the FFQ. All food items were categorized into 33 food groups that were constructed according to the Dutch food composition table 2011 [32]. Final food groups are described in Supplementary Table S1. Total intake of food groups was recalculated to relative intake (g/d per 1000 kcal) using total energy intake to simplify comparison of participants.

Principal component analysis was used to investigate data-driven dietary patterns among participants [33]. Food group data were log-transformed using the natural logarithm and Z-standardized before performing principal component analysis. As a result, the intake of all food groups has a mean of zero and a variance of one, which is important since the results of components highly depend on the variance of each variable [34]. In case a certain food group was not consumed, i.e. 0 g/day per 1000 kcal, 0.001 was added to the food group sum to allow log transformation. The number of dietary patterns was decided based on the components with eigenvalues  $> 1.0$ , the scree plot and the interpretability of the components [35]. The remaining components were orthogonally rotated for ease of interpretation and labels in

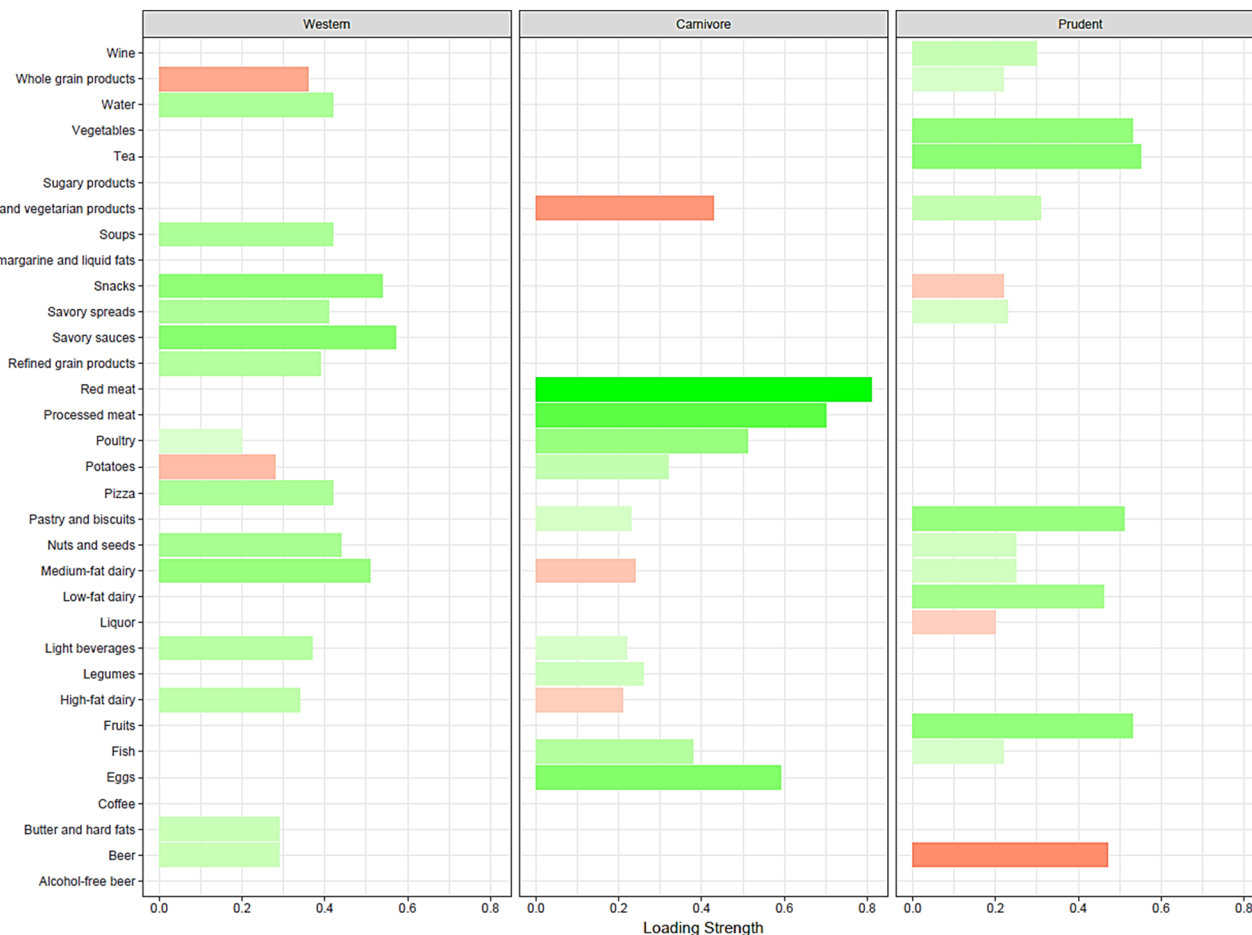
accordance with the included food groups were given. Positive and negative food group loadings  $>|0.2|$  were considered when naming the respective dietary pattern. Participants' scores were determined by multiplying the observed intake of all food groups by the factor loading for each of all the respective food groups [36].

Three dietary patterns were identified based on the available data, which were defined as a Western, Carnivore, and Prudent dietary pattern, see Fig. 1. The Western dietary pattern was characterized by a high intake of snacks, savory sauces and spreads, refined grains, pizza, high and medium-fat dairy, nuts and seeds, beer, and hard fats, and a low intake of whole grain products and potatoes. The Carnivore pattern was characterized by a high intake of red and processed meat, poultry, fish, eggs, and potatoes, and a low intake of soy and vegetarian products, and medium and high-fat dairy. Lastly, the Prudent pattern consisted of a high intake of vegetables, fruits, fish, nuts and seeds, low-fat dairy, tea, pastry and biscuits, and a low intake of beer.

## Biomarker analysis

Plasma samples were analyzed in four analytical batches at the International Agency for Research on Cancer (IARC) in Lyon, France. In total, 147 metabolites were measured using the Absolute<sup>IDQ</sup> p180 kit (Biocrates, Innsbruck, Austria). The analytical method [37, 38] characterizes up to 188 metabolites from five compound classes. Amino acids and biogenic amines were quantified (calibration curves, individual isotope-labelled internal standards) by ultra-high performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS). Lipids, including glycerophospholipids and sphingolipids, acylcarnitines, as well as the sum of hexose sugars were semi-quantified (one-point calibration, single representative internal standard) by flow injection analysis-tandem mass spectrometry (FIA-MS/MS).

Metabolites with  $> 20\%$  missing values ( $n = 13$ ), including values below the level of detection (LOD) and true missings, were removed from the dataset while remaining



**Fig. 1** Overview of food group loadings of the Western, Carnivore, and Prudent dietary patterns. Green and red bars represent positive and negative loading strengths, respectively. A more positive loading illustrates higher consumption of a specific food group, while a more

negative loading characterizes lower consumption of the food group. Only food group loadings  $>|0.2|$  were considered to contribute to the dietary pattern and visualized to improve readability

missing values were imputed in line with previous studies [38–40]. Briefly, missing values below the LOD were imputed by half of the batch-specific LOD and values below or above the quantitative range were replaced by the lower or upper limit of quantification, respectively. Subsequently, to normalize distributions, metabolite concentrations were log transformed using the natural logarithm and were Z-standardized to allow comparison of estimates among metabolites. In total, 134 metabolites were included in the current analysis (Supplementary Table S2), consisting of 12 acylcarnitines (Cx:y), 21 amino acids, 8 biogenic amines, 78 glycerophospholipids (10 lysophosphatidylcholines (lysoPC) and 68 phosphatidylcholines [PC diacyl (aa) and acyl-alkyl (ae)], 14 sphingolipids (SM Cx:y) and the sum of hexoses. The abbreviation Cx:y is used to describe the total number of carbons and double bonds in the alkyl chains, respectively.

### Statistical analysis

Clinical, demographic and lifestyle characteristics were described using descriptive analyses. Linear regression models were used to determine the associations between diet quality indices and dietary patterns as independent variables and concentrations of metabolites as dependent variables. Diet quality indices and dietary patterns were analyzed as continuous variables as well as analyzed in tertiles, for which the lowest tertile, corresponding to the lowest intake of the exposure, was used as the reference category. *P*-trend values were computed for tertiles using the medians of the corresponding tertiles.

All models were adjusted for age at diagnosis (continuous), sex, BMI (continuous), smoking status (current smoker/former smoker/never smoker), analytical batch (1–4), and cancer stage (stage I/stage II/ stage III/ stage IV). The basis for assessing whether covariates should be included in the final model were existing evidence, biological plausibility and whether the regression coefficient of interest changed by > 10% after adding the potential covariate.

Furthermore, to explore the consistency in the observed associations between dietary exposures and plasma metabolites, we evaluated the top-15 metabolites (based on the smallest *p*-value for trend over tertiles) associated with diet quality indices and dietary patterns using a heatmap.

Sensitivity analyses were conducted excluding patients from whom blood was collected during or after any type of treatment, i.e. (neo-) adjuvant chemotherapy and/or surgery (*n* = 19) and excluding stage IV patients (*n* = 7).

All statistical analyses were performed in R, version 3.4.0 and SAS, version 9.4. After correction for multiple testing, using false discovery rate (FDR) according to the Benjamin-Hochberg procedure [41, 42], a *p* value ( $p_{\text{FDR}}$ ) < 0.05 was considered statistically significant.

## Results

### Study population characteristics

Characteristics of the overall study population are summarized in Table 1. The mean (SD) age of the 195 colorectal cancer patients was 66 (9) years and almost 60% of the study population was male. Mean (SD) BMI was 25.6 (4.9) kg/m<sup>2</sup>, around ten percent were current smokers and participants had a stage I (27%), stage II (33%), stage III (36%) or stage IV (4%) cancer. Distal colon cancer was diagnosed in 36%

**Table 1** Baseline characteristics of the overall study population

	Study population
Number of participants	195
Male sex, <i>n</i> (%)	115 (59)
Age	66.2 ± 9.1
Body mass index (kg/m <sup>2</sup> )	25.6 ± 4.9
Underweight, <18.5, <i>n</i> (%)	3 (2)
Normal weight, 18.5–24.9, <i>n</i> (%)	85 (44)
Overweight, 25–29.9, <i>n</i> (%)	79 (41)
Obese, ≥30, <i>n</i> (%)	28 (14)
Smoking, <i>n</i> (%)	
Current	19 (10)
Former	118 (60)
Never	58 (30)
Stage, <i>n</i> (%)	
I	53 (27)
II	65 (33)
III	70 (36)
IV	7 (4)
Tumor site <sup>a</sup> , <i>n</i> (%)	
Colon—proximal	56 (29)
Colon—distal	71 (36)
Rectal	68 (35)
Treatment, <i>n</i> (%)	
Surgery	194 (99)
Neo-adjuvant treatment	60 (31)
Total energy intake (kcal/day)	1856 ± 559
Total WCRF dietary score <sup>b</sup>	2.1 ± 0.7
Total DHD15-index <sup>c</sup>	73.8 ± 14.1

Numbers are presented as mean ± SD or median (IQR) unless mentioned otherwise

<sup>a</sup>Proximal consisting of: hepatic flexure, transverse colon, cecum, appendix, ascending colon; distal consisting of: descending colon, sigmoid colon, splenic flexure; rectal consisting of: rectosigmoid junction, rectum

<sup>b</sup>Adherence to the dietary recommendations of the World Cancer Research Fund (WCRF), ranged 0 to 5 [18]

<sup>c</sup>Adherence to the Dutch Healthy Diet guidelines 2015 (DHD15), ranged 0 to 130 [19]



of the study population, proximal colon cancer and rectal cancer in 29% and 35%, respectively. The study population had a mean (SD) total energy intake of 1856 (559) kcal/day. The mean (SD) WCRF dietary score and DHD15-index were 2.1 (0.7) and 73.8 (14.1), respectively. There were no consistent different directions in diet quality indices and dietary patterns when comparing colorectal cancer stages (data not shown). Baseline characteristics of the participants in the lowest and highest tertile of each diet quality index and dietary pattern are shown in Supplementary Table S3.

### Diet quality indices

Table 2 presents the top-15 metabolites based on the smallest value of  $p_{\text{trend}}$  across tertiles for the analyses between diet quality indices and metabolite concentrations, which were ranked by the  $p_{\text{trend}}$  across tertiles of the diet quality indices. A higher concordance of the WCRF dietary score was statistically significantly associated after FDR adjustment with lower concentrations of ten phosphatidylcholines over increasing tertiles. Each one-point increase in the WCRF dietary score also showed statistically significantly lower concentrations of four of the above-mentioned ten phosphatidylcholines (PC ae C36:3, PC ae C36:4, PC aa C36:3, and PC aa C38:3).

The DHD15-index was not statistically significantly associated with plasma metabolites after FDR adjustment when analyzed by tertiles of the DHD15-index and continuously (Table 2). An overview of all the results on the association between the diet quality indices and all 134 metabolites is available in Supplementary Table S2.

### Dietary patterns

The top-15 metabolites resulting from analysis of associations between the Western, Carnivore, and Prudent pattern and plasma metabolites, ranked by the  $p_{\text{trend}}$  across tertiles of the dietary patterns, are shown in Table 3. No linear trend was observed over increasing tertiles of the Western pattern in relation to plasma metabolites. In contrast, every SD increase in consumption of the Western pattern was statistically significantly associated with 35 metabolites (Table 3 and Supplementary Table S2).

A linear trend was observed between increasing tertiles of the Carnivore pattern and higher concentrations of two phosphatidylcholines. Similarly, every SD increase in the intake of the Carnivore pattern was also statistically significantly associated with higher concentrations of PC aa C38:0 ( $p_{\text{FDR}}$ : 0.001) and PC ae 38:6 ( $p_{\text{FDR}}$ : 0.01).

The Prudent pattern was not statistically significantly associated with any metabolite when evaluating the linear trend, as well as when testing each SD increase in consumption of the Prudent pattern. Results of the analyses between

the dietary patterns and all 134 metabolites are provided in Supplementary Table S2.

### Overlap in the top-15 metabolites

Figure 2 illustrates the overlap in the observed top-15 metabolites (based on the smallest  $p$ -value for trend over tertiles) for each of the respective diet quality indices and dietary patterns. No overlap among acylcarnitines in the top-15 metabolites for each of the dietary exposures was observed. One amino acid was overlapping; the Carnivore and Prudent pattern both showed a positive association with plasma tryptophan. Sarcosine, a biogenic amine, was positively associated with both the DHD15-index and the Prudent pattern.

Several glycerophospholipids showed overlap between the investigated dietary exposures. An increasing adherence to the WCRF dietary score recommendations, as well as an increasing adherence to the DHD15-index recommendations were associated with decreasing concentrations of plasma phosphatidylcholine PC aa C32:1. Positive associations were observed between the DHD15-index and the Carnivore pattern and PC aa C38:6 concentrations. Inverse associations were observed for the WCRF dietary score, the DHD15-index and the Prudent pattern in relation to phosphatidylcholine PC aa C40:4.

Four phosphatidylcholines (PC ae 34:1, PC ae C34:2, PC ae C36:3, and PC ae C38:3) were among the top-15 metabolites associated with the WCRF dietary score and the Western pattern. Concentrations of these phosphatidylcholines were lower with increasing adherence to the WCRF dietary recommendations, while higher concentrations were observed with a higher consumption of the Western pattern. In addition, the WCRF dietary score and DHD15-index were inversely associated with two phosphatidylcholines (PC ae C36:4 and PC ae C38:5), while the opposite was observed for the Carnivore pattern. A higher DHD15-index and a higher intake of the Carnivore and Prudent pattern were associated with higher concentrations of PC ae C40:6.

Two sphingolipids overlapped between the dietary exposures. A higher consumption of the Western and Prudent pattern were both associated with a higher consumption of plasma SM (OH) C14:1. A higher DHD-15 index score reported higher concentrations of SM (OH) C22:2, and, in line, higher intakes of the Prudent pattern showed higher concentrations of SM (OH) C:22.

### Sensitivity analyses

Sensitivity analyses excluding patients from whom blood was collected during or after any type of treatment, i.e. (neo-) adjuvant chemotherapy and/or surgery ( $n = 19$ ) and excluding stage IV patients ( $n = 7$ ) showed similar beta coefficients between the dietary exposures and plasma

**Table 2** Top-15 plasma metabolites associated with diet quality indices, ranked by  $P_{trend}$

Metabolite	Continuous			Tertiles						
	1			2			3			
	$\beta$	95% CI	$P_{FDR}^c$	$\beta$	95% CI	$P_{FDR}^c$	$\beta$	95% CI	$P_{FDR}^c$	$P_{trend}^c$
<b>Diet quality index: WCRF dietary score<sup>a,b</sup></b>										
No. of participants	195			56			62			
Mean score (range)	2.1 (0.50–4.25)			2.10 (2.00–2.25)			2.80 (2.50–4.25)			
1	PC ae C36:3	-0.38 (-0.58;-0.19)	0.02	-0.37 (-0.69;-0.06)	0.39		-0.67 (-0.97;-0.36)	0.003		0.005
2	PC ae C36:4	-0.37 (-0.58;-0.16)	0.03	-0.29 (-0.63;0.04)	0.49		-0.67 (-0.99;-0.35)	0.004		0.006
3	PC aa C36:3	-0.34 (-0.54;-0.14)	0.04	-0.08 (-0.40;0.24)	0.91		-0.59 (-0.90;-0.28)	0.01		0.01
4	PC aa C38:3	-0.34 (-0.53;-0.15)	0.03	-0.11 (-0.42;0.20)	0.87		-0.53 (-0.82;-0.23)	0.01		0.02
5	PC ae C34:2	-0.30 (-0.50;-0.10)	0.08	-0.34 (-0.66;-0.01)	0.47		-0.55 (-0.86;-0.24)	0.01		0.02
6	PC ae C38:4	-0.29 (-0.49;-0.09)	0.09	-0.24 (-0.57;0.08)	0.61		-0.56 (-0.87;-0.25)	0.01		0.02
7	PC aa C40:4	-0.30 (-0.51;-0.09)	0.09	-0.26 (-0.60;0.08)	0.60		-0.52 (-0.85;-0.20)	0.03		0.04
8	PC ae C38:3	-0.27 (-0.46;-0.08)	0.09	-0.05 (-0.36;0.26)	0.92		-0.48 (-0.77;-0.18)	0.03		0.04
9	PC ae C38:5	-0.30 (-0.50;-0.10)	0.08	-0.22 (-0.55;0.11)	0.71		-0.52 (-0.83;-0.21)	0.03		0.04
10	PC ae C40:4	-0.24 (-0.44;-0.04)	0.15	-0.32 (-0.64;0.01)	0.49		-0.50 (-0.80;-0.19)	0.03		0.04
11	PC aa C34:2	-0.25 (-0.47;-0.04)	0.15	-0.12 (-0.47;0.23)	0.87		-0.46 (-0.80;-0.13)	0.08		0.07
12	PC aa C32:1	-0.26 (-0.47;-0.05)	0.15	-0.16 (-0.51;0.19)	0.82		-0.44 (-0.78;-0.11)	0.09		0.09
13	PC aa C34:1	-0.27 (-0.48;-0.06)	0.14	-0.12 (-0.46;0.23)	0.87		-0.44 (-0.77;-0.10)	0.09		0.09
14	PC ae C32:1	-0.24 (-0.45;-0.04)	0.15	-0.29 (-0.63;0.04)	0.49		-0.44 (-0.76;-0.12)	0.08		0.09
15	PC ae C34:1	-0.25 (-0.45;-0.06)	0.14	-0.19 (-0.51;0.13)	0.74		-0.41 (-0.72;-0.11)	0.08		0.09
<b>Diet quality index: DHD15-index<sup>a,d</sup></b>										
No. of participants	195			65			65			
Mean score (range)	73.8 (40.9–114.8)			73.5 (68.4–78.9)			89.2 (78.9–113.8)			
1	PC aa C40:4	-0.17 (-0.28;-0.07)	0.15	-0.40 (-0.75;-0.04)	0.30		-0.50 (-0.85;-0.14)	0.29		0.37
2	PC aa C42:0	0.14 (0.04;0.24)	0.20	0.11 (-0.22;0.44)	0.85		0.43 (0.09;0.76)	0.45		0.37
3	PC aa C42:1	0.13 (0.04;0.23)	0.20	0.26 (-0.07;0.59)	0.49		0.47 (0.14;0.81)	0.29		0.37
4	PC ae C36:4	-0.14 (-0.24;-0.03)	0.20	-0.59 (-0.94;-0.24)	0.15		-0.50 (-0.85;-0.15)	0.29		0.37
5	lysoPC a C17:0	0.05 (-0.06;0.16)	0.72	0.46 (0.11;0.81)	0.21		0.35 (0.00;0.71)	0.54		0.45
6	PC aa C32:1	-0.15 (-0.25;-0.04)	0.20	-0.40 (-0.76;-0.03)	0.33		-0.39 (-0.75;-0.02)	0.54		0.45
7	PC aa C38:6	0.10 (0.00;0.21)	0.36	-0.07 (-0.42;0.27)	0.88		0.35 (0.00;0.70)	0.54		0.45
8	PC aa C40:5	-0.14 (-0.24;-0.03)	0.20	-0.29 (-0.64;0.06)	0.48		-0.38 (-0.73;-0.03)	0.54		0.45
9	PC ae C38:5	-0.11 (-0.21;-0.01)	0.29	-0.47 (-0.81;-0.13)	0.21		-0.40 (-0.74;-0.06)	0.54		0.45
10	PC ae C40:6	0.10 (0.00;0.20)	0.36	0.04 (-0.29;0.38)	0.90		0.32 (-0.02;0.66)	0.54		0.45
11	PC ae C42:3	0.09 (-0.01;0.19)	0.42	-0.02 (-0.36;0.31)	0.96		0.35 (0.02;0.69)	0.54		0.45
12	PC ae C44:6	0.10 (0.00;0.20)	0.36	0.07 (-0.27;0.41)	0.88		0.32 (-0.02;0.66)	0.54		0.45

Table 2 (continued)

Metabolite	Continuous			Tertiles												
	1			2			3									
	$\beta$	95% CI	$P_{FDR}^c$	$\beta$	95% CI	$P_{FDR}^c$	$\beta$	95% CI	$P_{FDR}^c$	$P_{trend}^c$						
I3 Sarcosine	-0.10	(-0.21;0.00)	0.36	Ref			-0.05	(-0.41;0.31)	0.90			-0.37	(-0.73;-0.01)	0.54		0.45
I4 SM (OH) C22:2	0.05	(-0.04;0.14)	0.64	Ref			0.18	(-0.13;0.49)	0.56			0.28	(-0.03;0.59)	0.56		0.45
I5 C14:1	-0.12	(-0.23;-0.02)	0.26	Ref			-0.19	(-0.55;0.16)	0.61			-0.35	(-0.70;0.01)	0.54		0.45

<sup>a</sup>Tested using multiple linear regression models analyzing associations of diet quality index (continuous and in tertiles) as independent variable and log transformed Z-standardized metabolite concentrations as dependent variable. The continuous analysis is presented per one-point and ten-points increase for the WCRF dietary score and DHD15-index, respectively. Tertile cut-off scores were 1.75 and 2.25 for the WCRF dietary score and 68.3–78.9 for the DHD15-index. Regression models were adjusted for sex, age, analytical batch, body mass index (continuous), smoking status, and stage

<sup>b</sup>Adherence to the recommendations of the World Cancer Research Fund (WCRF), ranged 0 to 5

<sup>c</sup> $p$  value adjusted for multiple testing, using false discovery rate (FDR)

<sup>d</sup>Adherence to the Dutch Healthy Diet guidelines 2015 (DHD15), ranged 0 to 130

metabolite concentrations compared to the main analysis (data not shown).

## Discussion

The aim of the current study was to explore the associations between the diet, evaluated using diet quality indices and dietary patterns, and plasma metabolite concentrations in colorectal cancer patients. The WCRF dietary score and the Carnivore pattern were observed to be statistically significantly associated with several long-chain phosphatidylcholines. In addition, when exploring the overlap in the top-15 metabolites for the respective dietary exposures, several dietary exposures were associated with identical long-chain phosphatidylcholines, which strengthens the hypothesis that diet and plasma metabolite concentrations might be associated in colorectal cancer patients.

Better adherence to the WCRF dietary and DHD guidelines, reflecting a healthier diet, was, in general, associated with lower concentrations of phosphatidylcholines in colorectal cancer patients in this study. In contrast, higher intakes of the Western pattern, which is generally regarded as an unhealthier diet [22, 43–46], showed higher concentrations of phosphatidylcholines. Similarly, a higher intake of the Carnivore pattern was positively associated with phosphatidylcholines in the current study, suggesting that a higher intake of a diet with red and processed meat, poultry, fish, and eggs is associated with higher levels of phosphatidylcholines. Interestingly, a study by Schmidt et al. reported that a vegan diet was characterized by lower concentrations of phosphatidylcholines and sphingolipids compared to a diet high in animal products [12]. A previous study among healthy participants also reported decreased lipid concentrations, including lysophosphatidylcholines and other glycerophospholipids, after a two-month intervention assigning healthy individuals to a Mediterranean diet, which is generally low in animal products, except for fish, compared to a control diet. The control diet was based on the American Heart Association guidelines [47], which recommend to consume low-fat dairy products, fish, poultry, and lean meats regularly. In line with our results and the study of Schmidt et al. [12], this may suggest that a higher intake of animal products is associated with higher phosphatidylcholine concentrations, also among those with colorectal cancer. Further studies are needed to elucidate whether phosphatidylcholine metabolism may play a role in the colorectal cancer continuum.

Diet quality indices and dietary patterns have been linked to colorectal cancer survival previously [22, 48, 49]. Associations between dietary exposures and circulating metabolites in colorectal cancer patients may provide important leads for future research regarding the underlying mechanisms



**Table 3** Top-15 plasma metabolites associated with dietary patterns, ranked by  $P_{\text{trend}}$

Metabolite	Terriles																			
	Continuous					1					2					3				
	No. of participants	Mean score (range)	$\beta$	95% CI	$P_{\text{FDR}}^{\text{b,c}}$	Ref	$\beta$	95% CI	$P_{\text{FDR}}^{\text{b}}$	$\beta$	95% CI	$P_{\text{FDR}}^{\text{b}}$	$\beta$	95% CI	$P_{\text{FDR}}^{\text{b}}$	$P_{\text{trend}}^{\text{b}}$				
<b>Dietary pattern: western<sup>a</sup></b>																				
No. of participants	195					65	65	65	65	65	65	65	65	65	65	65				
Mean score (range)	0.0 (-5.2-2.4)					-1.0 (-5.2 to -0.3)	0.1 (-0.2-0.5)									0.9 (0.5-2.4)				
1	PC aa C28:1	0.31	(0.18;0.44)	<b>0.001</b>	Ref	0.18	(-0.14;0.50)	0.97	0.59	(0.26;0.92)	0.08	0.10	0.10			0.10				
2	PC aa C36:1	0.24	(0.10;0.38)	<b>0.01</b>	Ref	0.24	(-0.08;0.57)	0.97	0.54	(0.20;0.89)	0.09	0.10	0.10			0.10				
3	PC ae C30:0	0.20	(0.06;0.34)	<b>0.03</b>	Ref	0.27	(-0.07;0.60)	0.97	0.49	(0.15;0.84)	0.09	0.10	0.10			0.10				
4	PC ae C34:0	0.24	(0.10;0.39)	<b>0.01</b>	Ref	0.24	(-0.09;0.58)	0.97	0.50	(0.15;0.85)	0.09	0.10	0.10			0.10				
5	PC ae C36:1	0.25	(0.12;0.39)	<b>0.01</b>	Ref	0.25	(-0.07;0.57)	0.97	0.51	(0.18;0.85)	0.09	0.10	0.10			0.10				
6	PC ae C38:2	0.19	(0.05;0.33)	<b>0.04</b>	Ref	0.32	(-0.01;0.65)	0.97	0.47	(0.13;0.81)	0.09	0.10	0.10			0.10				
7	PC ae C38:3	0.19	(0.05;0.32)	<b>0.04</b>	Ref	0.23	(-0.08;0.54)	0.97	0.49	(0.16;0.81)	0.09	0.10	0.10			0.10				
8	PC ae C40:2	0.23	(0.10;0.37)	<b>0.01</b>	Ref	0.06	(-0.27;0.38)	0.97	0.50	(0.16;0.83)	0.09	0.10	0.10			0.10				
9	SM (OH) C14:1	0.22	(0.09;0.34)	<b>0.01</b>	Ref	0.11	(-0.18;0.41)	0.97	0.44	(0.13;0.75)	0.09	0.10	0.10			0.10				
10	SM (OH) C24:1	0.28	(0.14;0.42)	<b>0.01</b>	Ref	0.06	(-0.27;0.39)	0.97	0.52	(0.18;0.87)	0.09	0.10	0.10			0.10				
11	PC ae C36:3	0.23	(0.10;0.37)	<b>0.01</b>	Ref	0.27	(-0.05;0.60)	0.97	0.45	(0.11;0.79)	0.09	0.10	0.10			0.10				
12	PC aa C30:0	0.23	(0.09;0.38)	<b>0.02</b>	Ref	0.22	(-0.13;0.57)	0.97	0.47	(0.11;0.84)	0.10	0.11	0.11			0.11				
13	PC ae C34:1	0.23	(0.10;0.37)	<b>0.01</b>	Ref	0.25	(-0.07;0.57)	0.97	0.44	(0.11;0.78)	0.09	0.11	0.11			0.11				
14	PC ae C34:2	0.26	(0.12;0.40)	<b>0.01</b>	Ref	0.25	(-0.08;0.58)	0.97	0.44	(0.09;0.78)	0.10	0.11	0.11			0.11				
15	PC ae C36:0	0.25	(0.10;0.40)	<b>0.01</b>	Ref	0.26	(-0.09;0.61)	0.97	0.48	(0.12;0.84)	0.09	0.11	0.11			0.11				
<b>Dietary pattern: carnivore<sup>a</sup></b>																				
No. of participants	195					65	65	65	65	65	65	65	65	65	65	65				
Mean score (range)	0.0 (-5.2-1.5)					-0.9 (-5.2 to -0.1)	0.2 (-0.1-0.4)									0.7 (0.4-1.5)				
1	PC aa C38:0	0.32	(0.18;0.45)	<b>0.001</b>	Ref	0.31	(-0.01;0.64)	0.91	0.67	(0.33;1.01)	<b>0.01</b>	<b>0.02</b>	<b>0.02</b>			<b>0.02</b>				
2	PC ae C38:6	0.26	(0.12;0.39)	<b>0.01</b>	Ref	0.24	(-0.08;0.56)	0.91	0.68	(0.35;1.01)	<b>0.01</b>	<b>0.02</b>	<b>0.02</b>			<b>0.02</b>				
3	PC ae C36:5	0.18	(0.04;0.32)	0.12	Ref	0.19	(-0.14;0.51)	0.91	0.66	(0.32;0.99)	<b>0.01</b>	0.05	0.05			0.05				
4	PC aa C36:0	0.22	(0.09;0.36)	0.03	Ref	0.18	(-0.15;0.52)	0.91	0.59	(0.25;0.94)	<b>0.02</b>	0.08	0.08			0.08				
5	PC ae C38:5	0.20	(0.06;0.34)	0.05	Ref	0.24	(-0.09;0.57)	0.91	0.58	(0.24;0.92)	<b>0.02</b>	0.08	0.08			0.08				
6	PC aa C38:6	0.27	(0.13;0.40)	0.01	Ref	0.19	(-0.15;0.53)	0.91	0.51	(0.16;0.86)	0.06	0.10	0.10			0.10				
7	PC aa C40:6	0.29	(0.16;0.43)	0.00	Ref	0.10	(-0.23;0.43)	0.91	0.50	(0.16;0.84)	0.06	0.10	0.10			0.10				
8	PC aa C42:2	0.13	(-0.01;0.27)	0.35	Ref	0.22	(-0.11;0.56)	0.91	0.50	(0.15;0.84)	0.06	0.10	0.10			0.10				
9	PC ae C36:4	0.24	(0.10;0.38)	0.02	Ref	0.25	(-0.10;0.59)	0.91	0.52	(0.16;0.87)	0.06	0.10	0.10			0.10				
10	PC ae C40:6	0.26	(0.12;0.39)	0.01	Ref	0.19	(-0.13;0.51)	0.91	0.51	(0.18;0.85)	0.06	0.10	0.10			0.10				
11	Isoleucine	0.12	(-0.02;0.26)	0.35	Ref	0.23	(-0.10;0.57)	0.91	0.45	(0.10;0.80)	0.10	0.10	0.10			0.10				

**Table 3** (continued)

Metabolite	Continuous			Tertiles							
				1	2		3				
	$\beta$	95% CI	$P_{FDR}^{b,c}$	Ref	$\beta$	95% CI	$P_{FDR}^b$	$\beta$	95% CI	$P_{FDR}^b$	$P_{trend}^b$
<i>I2</i> Methionine	0.13	(-0.01;0.28)	0.33	Ref	0.23	(-0.11;0.57)	0.91	0.44	(0.09;0.79)	0.10	0.10
<i>I3</i> Threonine	0.18	(0.04;0.33)	0.12	Ref	0.34	(-0.01;0.68)	0.91	0.45	(0.09;0.80)	0.10	0.10
<i>I4</i> Tryptophan	0.18	(0.04;0.32)	0.12	Ref	0.35	(0.01;0.70)	0.91	0.49	(0.13;0.85)	0.08	0.10
<i>I5</i> Tyrosine	0.10	(-0.05;0.24)	0.56	Ref	0.30	(-0.05;0.65)	0.91	0.45	(0.09;0.81)	0.10	0.10
Dietary pattern: prudent <sup>a</sup>											
No. of participants	195			65	65			65			
Mean score (range)	0.0 (-3.1-1.8)			-1.1 (-3.1 to -0.3)	0.2 (-0.2-0.5)			1.0 (0.5-1.8)			
<i>I</i> PC aa C40:4	-0.22	(-0.39;-0.06)	0.11	Ref	-0.39	(-0.74;-0.04)	0.68	-0.54	(-0.9;-0.18)	0.09	0.10
<i>2</i> PC ae C30:2	0.20	(0.06;0.34)	0.10	Ref	0.42	(0.13;0.72)	0.68	0.50	(0.20;0.81)	0.09	0.10
<i>3</i> SM (OH) C14:1	0.20	(0.06;0.35)	0.10	Ref	0.27	(-0.03;0.58)	0.73	0.50	(0.18;0.81)	0.09	0.10
<i>4</i> SM (OH) C16:1	0.22	(0.08;0.37)	0.10	Ref	0.26	(-0.06;0.57)	0.73	0.51	(0.18;0.83)	0.09	0.10
<i>5</i> C16	-0.27	(-0.44;-0.11)	0.06	Ref	-0.39	(-0.74;-0.05)	0.68	-0.56	(-0.92;-0.20)	0.09	0.10
<i>6</i> SM (OH) C22:2	0.20	(0.06;0.34)	0.10	Ref	0.30	(0.00;0.6)	0.68	0.46	(0.15;0.77)	0.09	0.12
<i>7</i> C18:2	-0.28	(-0.45;-0.11)	0.06	Ref	-0.05	(-0.41;0.31)	0.91	-0.54	(-0.91;-0.17)	0.09	0.12
<i>8</i> C18:1	-0.28	(-0.45;-0.11)	0.06	Ref	-0.21	(-0.57;0.15)	0.91	-0.52	(-0.90;-0.15)	0.11	0.13
<i>9</i> SM (OH) C22:1	0.20	(0.05;0.35)	0.11	Ref	0.29	(-0.03;0.61)	0.73	0.43	(0.10;0.76)	0.17	0.19
<i>10</i> C18	-0.23	(-0.40;-0.07)	0.10	Ref	-0.32	(-0.67;0.03)	0.73	-0.47	(-0.84;-0.11)	0.17	0.19
<i>11</i> lysoPC a C16:1	-0.28	(-0.45;-0.11)	0.06	Ref	-0.35	(-0.71;0.01)	0.68	-0.44	(-0.82;-0.06)	0.21	0.20
<i>12</i> PC ae C40:6	0.13	(-0.03;0.29)	0.47	Ref	0.26	(-0.07;0.59)	0.73	0.42	(0.07;0.76)	0.19	0.20
<i>13</i> Aspartate	-0.12	(-0.30;0.05)	0.52	Ref	-0.06	(-0.42;0.30)	0.91	-0.46	(-0.83;-0.08)	0.19	0.20
<i>14</i> Tryptophan	0.18	(0.01;0.35)	0.25	Ref	0.13	(-0.23;0.48)	0.91	0.44	(0.08;0.81)	0.19	0.20
<i>15</i> Sarcosine	-0.18	(-0.35;-0.01)	0.25	Ref	-0.34	(-0.70;0.01)	0.68	-0.41	(-0.78;-0.04)	0.26	0.20

<sup>a</sup>Tested using multiple linear regression models analyzing associations of dietary patterns (continuous, using log transformed Z-standardized scores, and in tertiles) as independent variable and log transformed Z-standardized metabolite concentrations as dependent variable. The continuous analysis is presented per each increase in SD (equaling 1) for the Western, Carnivore, and Prudent pattern, respectively. Tertile cut-off scores were -0.3 and 0.5, -0.1 and 0.4, and -0.3 and 0.5 for the Western, Carnivore, and Prudent pattern, respectively. Regression models were adjusted for sex, age, analytical batch, body mass index (continuous), smoking status, and stage

<sup>b</sup>*p* value adjusted for multiple testing, using false discovery rate (FDR)

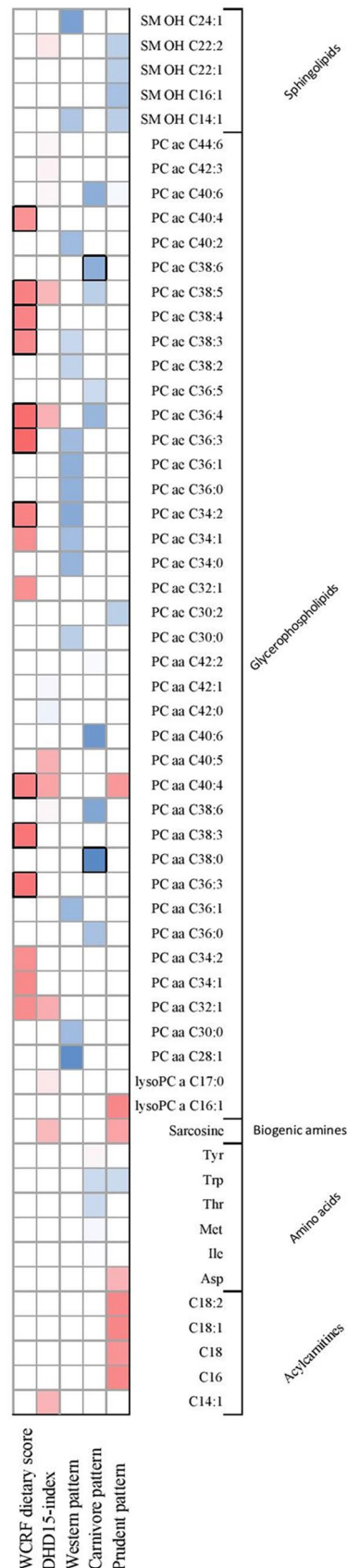
<sup>c</sup>In total, the Western pattern was statistically significantly associated (pFDR<0.05) with 35 plasma metabolites (see Supplementary Table S2), of which the top-15 metabolites were presented here

**Fig. 2** Heatmap illustrating the observed top-15 metabolites (based on the smallest *p*-value for trend over tertiles) associated with the diet quality indices, i.e. the WCRF dietary score and DHD15-index, and the dietary patterns, i.e. the Western, Carnivore, and Prudent pattern. The color is correlated to the observed  $\beta$  values; a darker blue color corresponds with a more positive association, while a darker red color represents a more inverse association between the dietary exposure and the plasma metabolite. Statistically significant associations are presented by a black box around the cell

between diet and colorectal cancer progression and survival. When these underlying mechanisms are identified, there is more solid scientific evidence to make nutritional recommendations for colorectal cancer survivors. However, since the current study is based on observational data only, it is not possible to clearly determine the causal relationships between dietary exposures and phosphatidylcholines in colorectal cancer patients. A previous study suggested that cancer cells display an elevated production of phosphatidylcholines, as part of enhanced lipogenesis in cancer cells [50], to further promote proliferation and evade apoptosis [51]. Given our findings that diet seems to be associated with phosphatidylcholines in colorectal cancer patients, and thus may theoretically support the hypothesized neoplastic growth, further studies studying phosphatidylcholines in relation to colorectal cancer recurrence and survival might be of interest.

The main strength of the current study is that this is, to the best of our knowledge, the first study investigating the associations between dietary exposures and plasma metabolites in a diseased population, i.e. colorectal cancer patients, using different approaches. When exploring the top-15 metabolites associated with the investigated dietary exposures, several phosphatidylcholines were observed to overlap between our exposures. This may indicate that the reported associations in our study population between dietary exposures and plasma metabolites are robust findings.

One of the limitations is that non-fasted blood samples were used for the current study and, as a result, we cannot rule out the possibility that some observed associations might be related to recent occasional dietary intake [38]. Our study was also limited to the metabolites included in the kit, while other metabolites might also be associated with the various dietary exposures. Following the presented results, a lipid-focused approach is of interest when investigating the association between diet and metabolites in colorectal cancer patients in the future. Lipid species, such as phosphatidylcholines, possess different physicochemical properties [52], and the methods used in the current and previous studies [11, 12] do not allow in-depth interpretation of the individual fatty acid compositions. Our relatively small sample size did not allow comparison of different colorectal cancer stages [40] and subtypes, although associations between diet and metabolites could potentially be related to specific tumor



characteristics [53]. Lastly, we were not able to analyze the potential associations between diet, metabolites and colorectal cancer recurrence or survival.

In summary, we reported that the WCRF dietary score and the Carnivore pattern are associated with plasma concentrations of phosphatidylcholines in colorectal cancer patients. Several phosphatidylcholines were also observed to overlap between the dietary exposures when comparing the top-15 metabolites. Our findings should be replicated in larger study populations to allow more in-depth analysis regarding colorectal cancer stages and subtypes to also explore the role of nutritional metabolites in the colorectal cancer continuum. Furthermore, future studies should investigate the association between nutritional metabolites and colorectal cancer recurrence and survival. These explorative analyses might provide additional information about the potential underlying mechanisms of dietary intake in colorectal cancer patients, and the potential relationship with recurrence and survival.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00394-021-02488-1>.

**Acknowledgements** The authors would like to thank all participants, the involved co-workers in the participating hospitals, and the COLON investigators at Wageningen University & Research.

**Authors' contributions** AGs, CMU, MPW, AS, and EK: acquisition of funding; AJMRG, DEK, EK, and FJBvD: contributed to the design and the conceptualization of this study; AJMRG, MvZ, FMK, JHWdW, and EW: contributed to the recruitment of participants and the data collection; PKR, DA, AGi and AS: performed metabolomics analyses, AJMRG: performed the statistical data analyses and drafted the manuscript; and all authors: critically read and revised the draft manuscript and read and approved the final manuscript.

**Funding** The COLON study was financially supported by Wereld Kanker Onderzoek Fonds (WKOF) and World Cancer Research Fund International (WCRF International) as well as by funds from grant 2014/1179 as part of the World Cancer Research Fund International Regular Grant Programme; Alpe d'Huzes/Dutch Cancer Society (UM 2012-5653, UW 2013-5927, UW 2015-7946); and ERA-NET on Translational Cancer Research (TRANSCAN: Dutch Cancer Society (UW2013-6397, UW2014-6877) and the Netherlands Organization for Health Research and Development (ZonMw), the Netherlands). The current study was also financially supported by Institut National du Cancer (Paris, grant number: 2014-007); Dr. Ulrich was supported by National Institutes of Health Grants U01 CA206110, R01 CA 189184, and R01 CA207371.

**Data availability** Because the data consist of identifying cohort information, some access restrictions apply, and therefore, they cannot be made publicly available. Data will be shared with permission, from the acting committee of the COLON Study. Requests for data can be sent to Dr. Fränzel van Duijnhoven, Division of Human Nutrition and Health, Wageningen University and Research, Netherlands (e-mail: [franzel.vanduijnhoven@wur.nl](mailto:franzel.vanduijnhoven@wur.nl)).

## Compliance with ethical standards

**Conflict of interest** Cornelia M. Ulrich (Director of the Comprehensive Cancer Center at Huntsman Cancer Institute, Salt Lake City, U.S.) oversees research funded by several pharmaceutical companies but has not received funding directly herself that could constitute a conflict to the current work. 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.

**Ethical standards** The study was performed in accordance with the Declaration of Helsinki. The COLON study was approved by the Committee of Research involving Human Subjects (region Arnhem-Nijmegen), the Netherlands. All participants included in the study provided a written informed consent.

**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/>.

## References

1. van Zutphen M, Kampman E, Giovannucci EL, van Duijnhoven FJB (2017) Lifestyle after colorectal cancer diagnosis in relation to survival and recurrence: a review of the literature. *Curr Colorect Cancer Rep* 13(5):370–401. <https://doi.org/10.1007/s11888-017-0386-1>
2. Menni C, Zhai G, MacGregor A, PrehnRmisch-Margl CW, Suhre K, Adamski J, Cassidy A, Illig T, Spector TD, Valdes AM (2013) Targeted metabolomics profiles are strongly correlated with nutritional patterns in women. *Metabolomics* 9(2):506–514. <https://doi.org/10.1007/s11306-012-0469-6>
3. O'Sullivan A, Gibney MJ, Brennan L (2010) Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies. *Am J Clin Nutr* 93(2):314–321
4. Brennan L (2018) Moving toward objective biomarkers of dietary intake. *The Journal of Nutrition* 148(6):821–822
5. Brennan L (2017) The nutritional metabolomics crossroads: how to ensure success for dietary biomarkers. *Am J Clin Nutr* 105(2):293–294
6. Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Rappaport SM, van der Hoof JJ, Wishart DS (2014) The food metabolome: a window over dietary exposure. *Am J Clin Nutr* 99(6):1286–1308
7. Ulaszewska MM, Weinert CH, Trimigno A, Portmann R, Andres Lacueva C, Badertscher R, Brennan L, Brunius C, Bub A, Capozzi F (2019) Nutrimetabolomics: an integrative action for

- metabolomic analyses in human nutritional studies. *Mol Nutr Food Res* 63(1):1800384
8. Playdon MC, Ziegler RG, Sampson JN, Stolzenberg-Solomon R, Thompson HJ, Irwin ML, Mayne ST, Hoover RN, Moore SC (2017) Nutritional metabolomics and breast cancer risk in a prospective study. *Am J Clin Nutr* 106(2):637–649
  9. AlEsa HB, Malik VS, Yuan C, Willett WC, Huang T, Hu FB, Tobias DK (2016) Dietary patterns and cardiometabolic and endocrine plasma biomarkers in US women, 2. *Am J Clin Nutr* 105(2):432–441
  10. Ahmad S, Moorthy M, Demler OV et al (2018) Assessment of risk factors and biomarkers associated with risk of cardiovascular disease among women consuming a Mediterranean diet. *JAMA Network Open* 1(8):e185708. <https://doi.org/10.1001/jamanetworkopen.2018.5708>
  11. Playdon MC, Moore SC, Derkach A, Reedy J, Subar AF, Sampson JN, Albanes D, Gu F, Kontto J, Lassale C (2017) Identifying biomarkers of dietary patterns by using metabolomics. *Am J Clin Nutr* 105(2):450–465. <https://doi.org/10.3945/ajcn.116.144501>
  12. Schmidt JA, Rinaldi S, Ferrari P, Carayol M, Achaintre D, Scalbert A (2015) Metabolic profiles of male meat eaters, fish eaters, vegetarians, and vegans from the EPIC-Oxford cohort. *Am J Clin Nutr*. <https://doi.org/10.3945/ajcn.115.111989>
  13. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Bittinger K, Chehoud C, Albenberg LG, Nessel L, Gilroy E (2016) Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. *Gut* 65(1):63–72
  14. Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* 13(1):3–9
  15. Romaguera D, Ward H, Wark PA, Vergnaud A-C, Peeters PH, van Gils CH, Ferrari P, Fedirko V, Jenab M, Boutron-Ruault M-C (2015) Pre-diagnostic concordance with the WCRF/AICR guidelines and survival in European colorectal cancer patients: a cohort study. *BMC Med* 13(1):107
  16. Inoue-Choi M, Robien K, Lazovich D (2013) Adherence to the WCRF/AICR guidelines for cancer prevention is associated with lower mortality among older female cancer survivors. *Cancer Epidemiol Prevent Biomark* 22(5):792–802
  17. Vergnaud A-C, Romaguera D, Peeters PH, Van Gils CH, Chan DS, Romieu I, Freisling H, Ferrari P, Clavel-Chapelon F, Fagherazzi G (2013) Adherence to the World Cancer Research Fund/American Institute for Cancer Research guidelines and risk of death in Europe: results from the European Prospective Investigation into Nutrition and Cancer cohort study. *Am Clin Nutr* 97(5):1107–1120
  18. Shams-White MM, Brockton NT, Mitrou P, Romaguera D, Brown S, Bender A, Kahle LL, Reedy J (2019) Operationalizing the 2018 world cancer research fund/American institute for cancer research (WCRF/AICR) cancer prevention recommendations: a standardized scoring system. *Nutrients* 11(7):1572
  19. Looman M, Feskens EJ, de Rijk M, Meijboom S, Biesbroek S, Temme EH, de Vries J, Geelen A (2017) Development and evaluation of the Dutch healthy diet index 2015. *Public Health Nutr* 20(13):2289–2299
  20. Van Lee L, Geelen A, Kieft-de Jong J, Witteman J, Hofman A, Vonk N, Jankovic N, Van Huysduyven EH, De Vries J, Van't Veer P (2016) Adherence to the Dutch dietary guidelines is inversely associated with 20-year mortality in a large prospective cohort study. *Eur J Clin Nutr* 70(2):262
  21. Franses HP, Beulens JW, May AM, Struijk EA, Boer JM, de Wit GA, Onland-Moret NC, van der Schouw YT, Bueno-de-Mesquita HB, Hoekstra J (2015) Dietary patterns in relation to quality-adjusted life years in the EPIC-NL cohort. *Prev Med* 77:119–124
  22. Meyerhardt JA, Niedzwiecki D, Hollis D, Saltz LB, Hu FB, Mayer RJ, Nelson H, Whittom R, Hantel A, Thomas J (2007) Association of dietary patterns with cancer recurrence and survival in patients with stage III colon cancer. *JAMA* 298(7):754–764
  23. van Ommen B, Keijzer J, Heil SG, Kaput J (2009) Challenging homeostasis to define biomarkers for nutrition related health. *Mol Nutr Food Res* 53(7):795–804. <https://doi.org/10.1002/mnfr.200800390>
  24. Gibney MJ, Walsh M, Brennan L, Roche HM, German B, van Ommen B (2005) Metabolomics in human nutrition: opportunities and challenges. *Am J Clin Nutr* 82(3):497–503. <https://doi.org/10.1093/ajcn/82.3.497>
  25. Guasch-Ferré M, Bhupathiraju SN, Hu FB (2018) Use of metabolomics in improving assessment of dietary intake. *Clin Chem* 64(1):82–98. <https://doi.org/10.1373/clinchem.2017.272344>
  26. Winkels RM, Heine-Bröring RC, Van Zutphen M, van Harten-Gerritsen S, Kok DE, Van Duynhoven FJ, Kampman E (2014) The COLON study: colorectal cancer: longitudinal, observational study on nutritional and lifestyle factors that may influence colorectal tumour recurrence, survival and quality of life. *BMC Cancer* 14(1):1
  27. Siebelink E, Geelen A, de Vries JH (2011) Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults. *Br J Nutr* 106(2):274–281
  28. Streppel MT, de Vries JH, Meijboom S, Beekman M, de Craen AJ, Slagboom PE, Feskens EJ (2013) Relative validity of the food frequency questionnaire used to assess dietary intake in the Leiden longevity study. *Nutr J* 12(1):75
  29. Van Leersum N, Sniijders H, Henneman D, Kolfshoten N, Gooiker G, Ten Berge M, Eddes E, Wouters M, Tollenaar R, Bemelman W (2013) The Dutch surgical colorectal audit. *Eur J Surg Oncol (EJSO)* 39(10):1063–1070
  30. Kromhout D, Spaaij C, de Goede J, Weggemans R (2016) The 2015 Dutch food-based dietary guidelines. *Eur J Clin Nutr* 70(8):869
  31. van Zutphen M, Boshuizen HC, Kok DE, van Baar H, Geijzen AJ, Wesselink E, Winkels RM, van Halteren HK, de Wilt JH, Kampman E (2019) Colorectal cancer survivors only marginally change their overall lifestyle in the first 2 years following diagnosis. *J Cancer Surviv* 13:956–967
  32. Netherlands Nutrition Center: NEVO Nederlandse Voedingsmiddelen Tabel 2011 (In English: Dutch Food Composition Table). <http://nevo-online.rivm.nl/>. Accessed June 2018
  33. Borges CA, Rinaldi AE, Conde WL, Mainardi GM, Behar D, Slater B (2015) Dietary patterns: a literature review of the methodological characteristics of the main step of the multivariate analyses. *Rev Br Epidemiol* 18(4):837–857. <https://doi.org/10.1590/1980-5497201500040013>
  34. Reimann C, Filzmoser P, Garrett RG, Dutter R (2008) Statistical data analysis explained: applied environmental statistics with R. Principal component analysis (PCA) and factor analysis (FA).
  35. Franses HP, May AM, Stricker MD, Boer JM, Hennig C, Rosseel Y, Ocke MC, Peeters PH, Beulens JW (2014) A posteriori dietary patterns: how many patterns to retain? *J Nutr* 144(8):1274–1282. <https://doi.org/10.3945/jn.113.188680>
  36. McCann SE, Marshall JR, Brasure JR, Graham S, Freudenheim JL (2001) Analysis of patterns of food intake in nutritional epidemiology: food classification in principal components analysis and the subsequent impact on estimates for endometrial cancer. *Public Health Nutrition* 4(5):989–997
  37. Siskos AP, Jain P, Römisch-Margl W, Bennett M, Achaintre D, Asad Y, Marney L, Richardson L, Koulman A, Griffin JL (2016) Interlaboratory reproducibility of a targeted metabolomics platform for analysis of human serum and plasma. *Anal Chem* 89(1):656–665
  38. Carayol M, Licaj I, Achaintre D, Sacerdote C, Vineis P, Key TJ, Moret NCO, Scalbert A, Rinaldi S, Ferrari P (2015) Reliability of serum metabolites over a two-year period: a targeted



- metabolomic approach in fasting and non-fasting samples from EPIC. *PLoS ONE* 10(8):e0135437
39. van Roekel EH, Trijsburg L, Assi N, Carayol M, Achaintre D, Murphy N, Rinaldi S, Schmidt JA, Stepien M, Kaaks R (2018) Circulating metabolites associated with alcohol intake in the European prospective investigation into cancer and nutrition cohort. *Nutrients* 10(5):654
  40. Geijssen AJ, van Roekel EH, van Duijnhoven FJ, Achaintre D, Bachleitner-Hofmann T, Baierl A, Bergmann MM, Boehm J, Bours MJ, Brenner H (2019) Plasma metabolites associated with colorectal cancer stage: findings from an international consortium. *Int J Cancer* 146(12):3256–3266
  41. Cambiaghi A, Ferrario M, Masseroli M (2016) Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration. *Brief Bioinform*. <https://doi.org/10.1093/bib/bbw031>
  42. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B (Methodol)* 57(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
  43. Fung T, Hu FB, Fuchs C, Giovannucci E, Hunter DJ, Stampfer MJ, Colditz GA, Willett WC (2003) Major dietary patterns and the risk of colorectal cancer in women. *Arch Intern Med* 163(3):309–314
  44. Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, Sampson L, Willett WC (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* 69(2):243–249
  45. Kwan ML, Weltzien E, Kushi LH, Castillo A, Slattery ML, Caan BJ (2008) Dietary patterns and breast cancer recurrence and survival among women with early-stage breast cancer. *J Clin Oncol* 27(6):919–926
  46. Varraso R, Garcia-Aymerich J, Monier F, Le Moual N, De Batlle J, Miranda G, Pison C, Romieu I, Kauffmann F, Maccario J (2012) Assessment of dietary patterns in nutritional epidemiology: principal component analysis compared with confirmatory factor analysis. *Am J Clin Nutr* 96(5):1079–1092. <https://doi.org/10.3945/ajcn.112.038109>
  47. Bondia-Pons I, Martinez JA, de la Iglesia R, Lopez-Legarrea P, Poutanen K, Hanhineva K, MdlÁ Z (2015) Effects of short-and long-term Mediterranean-based dietary treatment on plasma LC-QTOF/MS metabolic profiling of subjects with metabolic syndrome features: the metabolic syndrome reduction in Navarra (RESMENA) randomized controlled trial. *Mol Nutr Food Res* 59(4):711–728
  48. Fung TT, Kashambwa R, Sato K, Chiuev SE, Fuchs CS, Wu K, Giovannucci E, Ogino S, Hu FB, Meyerhardt JA (2014) Post diagnosis diet quality and colorectal cancer survival in women. *PLoS ONE* 9(12):e115377
  49. Ratjen I, Schafmayer C, di Giuseppe R, Waniek S, Plachta-Danielzik S, Koch M, Nothlings U, Hampe J, Schlesinger S, Lieb W (2017) Postdiagnostic Mediterranean and healthy Nordic dietary patterns are inversely associated with all-cause mortality in long-term colorectal cancer survivors. *J Nutr* 147(4):636–644. <https://doi.org/10.3945/jn.116.244129>
  50. Koundouros N, Pouligiannis G (2020) Reprogramming of fatty acid metabolism in cancer. *Br J Cancer* 122(1):4–22. <https://doi.org/10.1038/s41416-019-0650-z>
  51. Ridgway ND (2013) The role of phosphatidylcholine and choline metabolites to cell proliferation and survival. *Crit Rev Biochem Mol Biol* 48(1):20–38. <https://doi.org/10.3109/10409238.2012.735643>
  52. German JB, Gillies LA, Smilowitz JT, Zivkovic AM, Watkins SM (2007) Lipidomics and lipid profiling in metabolomics. *Curr Opin Lipidol* 18(1):66–71
  53. McGee EE, Kiblawi R, Playdon MC, Eliassen AH (2019) Nutritional metabolomics in cancer epidemiology: current trends, challenges, and future directions. *Curr Nut Rep* 8:187–201