Evaluation of dietary patterns among Norwegian postmenopausal women using plasma carotenoids as biomarkers

Marianne S. Markussen[1], Marit B. Veierød[1,2], Amrit K. Sakhi[3], Merete Ellingjord-Dale[4], Rune Blomhoff[1,5], Giske Ursin[1,4,6]* and Lene F. Andersen[1]

[1]*Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway*
[2]*Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway*
[3]*Norwegian Institute of Public Health, Oslo, Norway*
[4]*Cancer Registry of Norway, Oslo, Norway*
[5]*Division of Cancer Medicine, Surgery and Transplantation, Oslo University Hospital, Oslo, Norway*
[6]*Department of Preventive Medicine, University of Southern California, Los Angeles, CA, USA*

Abstract

A number of studies have examined dietary patterns in various populations. However, to study to what extent such patterns capture meaningful differences in consumption of foods is of interest. In the present study, we identified important dietary patterns in Norwegian postmenopausal women (age 50–69 years, n 361), and evaluated these patterns by examining their associations with plasma carotenoids. Diet was assessed by a 253-item FFQ. These 253 food items were categorised into forty-six food groups, and dietary patterns were identified using principal component analysis. We used the partial correlation coefficient (r_{adj}) and multiple linear regression analysis to examine the associations between the dietary patterns and the plasma carotenoids α-carotene, β-carotene, β-cryptoxanthin, lutein, lycopene and zeaxanthin. Overall, four dietary patterns were identified: the 'Western'; 'Vegetarian'; 'Continental'; 'High-protein'. The 'Western' dietary pattern scores were significantly inversely correlated with plasma lutein, zeaxanthin, lycopene and total carotenoids ($-0.25 \leq r_{adj} \leq -0.13$). The 'Vegetarian' dietary pattern scores were significantly positively correlated with all the plasma carotenoids ($0.15 \leq r_{adj} \leq 0.24$). The 'Continental' dietary pattern scores were significantly inversely correlated with plasma lutein and α-carotene ($r_{adj} = -0.13$). No significant association between the 'High-protein' dietary pattern scores and the plasma carotenoids was found. In conclusion, the healthy dietary pattern, the 'Vegetarian' pattern, is associated with a more favourable profile of the plasma carotenoids than our unhealthy dietary patterns, the 'Western' and 'Continental' patterns.

Key words: Dietary patterns: Principal component analysis: Plasma carotenoids: FFQ: Nutritional biomarkers: Postmenopausal women: Validation

Our diet has an important role in promoting health and preventing disease. Traditionally, nutritional epidemiology has focused on the intake of single food items or single nutrients, and not on the entire diet[1]. However, there is a complex relationship between diet and disease. Since we consume foods and nutrients in different combinations, analysing food consumption as dietary patterns is a complementary approach to understand the diet–disease relationship[2]. There is a unique dietary culture in every country that is reflected in the dietary patterns, and few studies have been conducted on empirically derived food patterns in adult Norwegian women[3–5]. Use of nutritional biomarkers in the evaluation of dietary patterns could provide knowledge on whether the derived patterns are meaningful. A number of studies have investigated the associations between dietary patterns and nutritional biomarkers[6–16]. A dietary pattern that is interpreted as a healthy dietary pattern has been reported to be positively associated with serum carotenoids, vitamin C and folate[6,7,9,10,13]. Carotenoids in blood are classified into concentration biomarkers and have been widely used as nutritional biomarkers for fruit and vegetable intake[17]. Of the approximately 600 carotenoids found in plant species[18], only α-carotene, β-carotene, β-cryptoxanthin, lutein, lycopene and zeaxanthin have been found in appreciable concentrations in human plasma[19]. Different populations tend to have different carotenoid profiles depending on their fruit and vegetable intake and lifestyle factors such as smoking status and BMI[20–22].

The aims of the present study were to (1) identify important dietary patterns in Norwegian postmenopausal women and (2) examine the association between the dietary patterns and

---

plasma carotenoid concentrations. We hypothesised that a healthy dietary pattern is associated with a more favourable profile of the plasma carotenoids than an unhealthy dietary pattern.

Methods

Study sample

The subjects in the present study were women participating in the Norwegian Breast Cancer Screening Program (NBCSP), a government-funded national screening programme administered by the Cancer Registry of Norway[23]. All Norwegian women aged 50–69 years were invited to a bilateral two-view mammogram biennially. The participation rate in the NBCSP was 77 %[24], with about 250 000 women being invited per year. In 2006/2007, the NBCSP invitation letter for mammographic screening included a question on willingness to complete a dietary questionnaire. A total of 67 527 women agreed to participate. In 2008, a consent form and a FFQ were sent to a random sample of 10 000 of these women. A total of 6974 returned the FFQ; of these, forty-six were not filled in and, therefore, excluded from the study. We further excluded 158 women who had missing data on height and weight, five women who had missing data on age and forty-one women who had missing data on smoking status. Furthermore, twenty women were excluded because of improbable height and weight status (seven with self-reported height <125 cm and thirteen with self-reported weight <30 or >170 kg). We excluded fifteen women because their age was not in the range of 50–69 years, four women because of BMI <15 or >50 kg/m² ; and 204 women because of implausible energy intake (<2100 or $>15 000$ kJ/d). This left us with a total sample of 6481 women (NBCSP subsample). Of the initial sample of 6974 women, more than 90 % agreed to provide saliva and blood samples. A self-collection kit containing necessary equipment for collecting both saliva and fingertip blood samples on a filter paper was sent to 4597 women. We received 3258 saliva and 3263 blood samples. Funding was available for a subset of laboratory analyses. The main focus of the parent study was breast cancer and mammographic density[25]; therefore, analyses of blood samples were restricted to women who also had an analogue mammogram (n 632) and who fulfilled the inclusion criteria for that study (n 387). In addition, we restricted analyses to women who followed the instructions given for the storage of blood samples. We were then left with blood samples from a subsample of 361 women in whom the carotenoid analyses were conducted. Thus, we had both dietary and plasma carotenoid data from these 361 women. The present study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human subjects were approved by the Regional Ethics Committee and the Norwegian Data Inspectorate. Written informed consent was obtained from all subjects.

Dietary assessment

A sixteen-page, 253-item FFQ was designed to capture the habitual food intake among Norwegian adults the preceding year, with an extra focus on the intake of fruit, vegetables, antioxidant-rich foods and beverages. It was based on a previously validated 180-item FFQ designed to cover the total energy intake in the Norwegian population[26], which was later expanded to a 270-item FFQ to cover the most antioxidant-rich foods and beverages in Norway[27]. Energy and food intake estimated from the 270-item FFQ was validated. Energy intake was evaluated against independent measures of energy expenditure using the ActiReg® system (PreMed AS) (motion detection), whereas 7 d weighed food records were used to study the relative validity of food and nutrient intake[27,28]. The correlation coefficient between energy intake and energy expenditure was 0·54. Correlation coefficients between the FFQ and the weighed food records were 0·41 for berries, 0·61 for fruit and 0·38 for vegetables[27]. This FFQ was also validated for ranking individuals according to their usual intake of fruit, juices and vegetables by using the method of triads with two independent and specific biomarkers of fruit and vegetables, the FFQ and 7 d weighed food records. The validity coefficients ranged from 0·60 to 0·94[28]. We revised the 270-item FFQ, by removing seventeen items that were seldom or never eaten, resulting in a 253-item FFQ. For each food item, participants indicated their frequency of consumption ranging from never/seldom to several times per d. The portion size per consumption was asked in household units such as slices, glasses, cups, pieces, spoons and teaspoons. When a respondent only answered about the frequency of consumption, but not the portion size, the food item was given the smallest portion size. If only the amount of the food item was given, the frequency was set to 0. The questionnaire also collected information about dietary supplements, age, height, weight, smoking status, physical activity, illnesses and medication. Daily intake of energy, nutrients and foods was computed using the food database AE-07 and KBS software system (version 4.9 2008; KBS) developed at the Department of Nutrition, University of Oslo, Norway. The food database AE-07 is based on the 2006 edition of the Norwegian food composition table (http://www.norwegianfoodcomp.no) supplemented with data from other tables and calculated recipes. Intakes from supplements were included in the calculations.

The 253 food items were categorised into forty-six food groups based on similarity in ingredients, nutrient profile or culinary usage (Table 1).

Blood collection and carotenoid analyses

A self-collection kit containing necessary equipment and a detailed instruction on how to collect fingertip blood samples on a filter paper were sent to the participants by mail. The first two spots on the filter paper were impregnated with a proprietary stabilising solution (Vitas AS). The blood samples were collected after fasting overnight, and each participant collected capillary blood sample from a fingertip directly on the filter paper (Whatman 903 paper; GE Healthcare), which subsequently dried for 8 h. The filter paper with the dried blood spots (DBS cards) was stored in an airtight alumina bag together with a silica drying medium (Whatman) and mailed

by regular mail to the study centre. The DBS cards were stored at −80°C at the study centre. Quantification of carotenoids in the DBS was performed by the contract laboratory Vitas AS. In short, five punches of 3·2 mm from each DBS were punched into vials, distilled water was added and mixed, proteins were precipitated, and carotenoids extracted with isopropanol as an internal standard (β-apo-8-carotenal; Sigma-Aldrich). HPLC−UV analysis was performed on an

Table 1. Food groupings used in the dietary pattern analysis

Food groups	Food items
Artificially sweetened beverages	Artificially sweetened soft drinks, artificially sweetened squash, artificially sweetened ice tea
Beer	Malt beer, pilsner, light beer
Butter	Butter, melted butter
Cakes and desserts	Dessert mousse, vanilla sauce, canned fruits, ice pop, muffins, chocolate cake, sponge cake with cream (with or without marzipan), sweet biscuits, Danish pastry, cinnamon bun/sweet bun with vanilla cream
Cheese	Regular and low-fat Norwegian brown cheese, regular and low-fat hard cheese, regular and low-fat cheese spread, cream cheese
Coffee	Coffee, boiled/press, instant, espresso, café latte, cappuccino
Eggs	Eggs
Fat-rich potatoes	Potato salad, potato gratin with cream, fried potatoes, home-made French fries, restaurant-made French fries
Fish	Caviar spread, roe paste, mackerel in tomato sauce, smoked salmon/brown trout, fried salmon/brown trout, sardine, herring, shrimps/crabs, fish cakes/fish pudding, fish balls, fish fingers, boiled cod/saithe/haddock/catfish/redfish, fried cod/saithe/haddock/catfish/redfish, smoked mackerel, fried mackerel, fish gratin, wok with seafood and vegetables
Fruit juices	Blueberry and aronia juice, cranberry and raspberry juice, rose hip and orange juice, orange juice, apple juice
Fruits and berries	Apple, pear, banana, orange, clementine, peach/nectarine, kiwi, grapes, melon, pomegranate, fruit as spreads, fresh fruit salad, prune, raisins, other dried fruits, blackberry, blueberry, raspberry, strawberry, cherry, cloudberry, rose hips, redcurrant, blackcurrant, cowberry
High-fat dairy products	High-fat milk, flavoured milk, sour cream, ice cream, whipped cream, high-fat yogurt
Herbs/spices	Dried and fresh basil, dried and fresh chilli, dried and fresh oregano, dried and fresh thyme, cinnamon, taco seasoning, grill seasoning, cardamom, curry powder, black pepper powder, sweet red pepper powder, dried rosemary, garlic, fresh dill, fresh ginger, fresh peppermint, fresh parsley
Legumes	Legumes
Liquor	Spirits, cider, cocktail
Low-fat dairy products	Skimmed and semi-skimmed milk, cultured/probiotic low-fat milk, low-fat yogurt drink, low-fat yogurt
Margarine	Normal and low-fat margarine
Mayonnaise	Regular and low-fat remoulade/mayonnaise, regular and low-fat spread with mayonnaise (Italian, shrimp, etc.)
Meat dishes	Mutton and cabbage stew, stew with meat, vegetables and potatoes, wok with meat and vegetables
Mustard	Mustard
Nuts and seeds	Cashew nuts, peanuts, peanut butter, walnuts, hazelnuts, almonds, pecan nuts, pine nuts, pistachio nuts, sesame seeds, sunflower seeds
Pasta	Noodles, pasta, pasta with tomato sauce
Pizza	Pizza
Potatoes	Boiled, mashed
Poultry	Grilled chicken, chicken fillet, chicken/turkey sausage, chicken/turkey grilled/wiener sausage
Processed meat	Meat balls, minced meat sauce, taco, kebab, lasagne, grilled/wiener sausage, minced meat sausage, bacon, pork chops, regular and low-fat liver paste, regular and low-fat saveloy, salami
Read meat and game	Beef, roast of lamb/beef/pork, roast of game, hamburger
Refined grains	White bread, crisp bread (wheat flour), hot dog bun, sweet muesli/breakfast cereal, pancake, rice porridge, waffle, sweet bun
Rice	Rice
Salad dressings	Thousand island dressing, oil and vinegar dressing, pesto
Salty snacks	Potato chips, other salty snacks
Sauce	Brown/white sauce, béarnaise/hollandaise sauce
Soup	Tomato soup
Soya sauce	Soya sauce
Sugar	Pure sugar
Sugar-sweetened beverages	Fruit juices with added sugar, squash with sugar, sugar-sweetened soft drinks, iced tea with sugar
Sweets	Chocolate, dark chocolate, extra dark chocolate, sweets/jelly sweets, sweet pastilles/candy
Sweeteners	Sweetener, sugar-free pastille
Sweet spreads	Regular and low-sugar jam, honey, chocolate/nut spread, other sweet spreads
Tea	Black tea, green tea, herb tea
Tomato sauce	Ketchup, tomato sauce, taco sauce
Vegetables	Carrot, cabbage, swede, cauliflower, broccoli, Brussels sprout, onion, spinach, sweet pepper, avocado, tomato, maize, frozen vegetables, mixed salad (with lettuce, cucumber, tomato and sweet pepper), vegetables as spread
Vegetarian food	Vegetarian food
Water	Tap water, mineral water
Wine	Red wine, white wine
Whole grains	Semi- and whole-grain bread, crisp bread (whole meal), oatmeal porridge and cereal, unsweetened muesli/breakfast cereal

1100-series HPLC system with a 1260 diode array detector (453 nm; Agilent Technologies). Separation was performed on a 3 μm YMC C30 column (150 mm × 4·6 mm inner diameter; YMC). Calibration was performed by analysing DBS calibrators spotted with full blood with known concentrations of carotenoids: lutein; zeaxanthin; β-cryptoxanthin; α-carotene; β-carotene; lycopene. The known concentrations are obtained by the analysis of serum from the same full blood. The calibrator for these values was NIST SRM-1950 (National Institute of Standards and Technology). A fixed haematocrit value of 2 was used to convert from DBS to plasma values. Quality control samples of DBS were run alongside the study samples. The total variation (CV%, n 50) for these quality control samples ranged from 7·8 to 9·0 % for lutein, β-carotene, lycopene and β-cryptoxanthin; 21·0 % for the low-abundant zeaxanthin; and 20·1 % for α-carotene. The total carotenoid concentration in plasma was calculated as the sum of the individual carotenoid concentrations quantified in this analysis.

Statistical analyses

Descriptive statistics were used to summarise the characteristics of the study sample (n 361) and the NBCSP subsample (without the study sample, n 6120). We studied the differences between the study sample and the NBCSP subsample by independent-samples t test for continuous variables and Pearson's χ^2 test for categorical variables.

Principal component analysis (PCA) was used to derive the dietary patterns based on the correlation matrix of the forty-six food groups. Before extracting the components, the suitability for using PCA was assessed by the Kaiser–Meyer–Olkin measure of sampling adequacy and Bartlett's test of sphericity[29]. The Kaiser–Meyer–Olkin value was 0·63, which is above the suggested minimum of 0·50[30], and Bartlett's test of sphericity was significant ($P < 0.001$), supporting the suitability of the data for PCA. To determine the number of meaningful components to retain, we considered the eigenvalue-one criterion, the scree test, the proportion of the variance accounted for and the interpretability criterion[1]. For interpretation purposes, varimax rotation was performed on the retained components. We considered food groups with a factor loading whose absolute value was ≥ 0.3 to load on that component. Finally, each woman's score was calculated for each of the retained components. We interpreted the retained components as dietary patterns and labelled them according to the more or less healthy combinations of food groups.

Table 2. Selected characteristics and plasma carotenoid concentrations in the study sample and in the Norwegian Breast Cancer Screening Program (NBCSP) subsample

(Medians, means, standard deviations and ranges; number of subjects and percentages)

	Study sample (n 361)				NBCSP subsample not selected for the present substudy* (n 6120)				
	Median	Mean	SD	Range	Median	Mean	SD	Range	P
Age (years)	57	57·1	4·1	50–65	58	57·9	4·5	50–69	<0·001
BMI (kg/m²)	24·7	25·5	4·3	15·0–45·2	24·9	25·6	4·0	15·4–49·5	0·71
Alcohol intake (g/d)	5·0	7·4	7·5	0–45·5	4·7	7·6	9·0	0–126·8	0·70
Energy intake (kJ/d)	8358	8621	2098	3547–14 741	8462	8697	2227	2210–14 975	0·53
Smokers									0·005
n		50				1215			
%		14				20			
Multivitamin users†									0·87
n		83				1395			
%		23				23			
Education‡									0·01
Primary and secondary school									
n		55				1230			
%		15				20			
Upper secondary school									
n		130				2422			
%		36				40			
Academy/college/university ≤4 years									
n		104				1443			
%		29				24			
Academy/college/university >4 years									
n		67				937			
%		19				15			
Plasma carotenoids (μmol/l)									
Lutein	0·198	0·232	0·134		–	–			
Zeaxanthin	0·044	0·047	0·021		–	–			
β-Cryptoxanthin	0·138	0·160	0·115		–	–			
α-Carotene	0·111	0·134	0·096		–	–			
β-Carotene	0·383	0·439	0·278		–	–			
Lycopene	0·237	0·257	0·125		–	–			
Total carotenoids	1·188	1·269	0·539		–	–			

* The NBCSP subsample without the study sample (n 6481 − 361 = 6120).
† n 323 for the study sample and n 5340 for the NBCSP subsample not selected for the present substudy.
‡ n 361 for the study sample and n 6119 for the NBCSP subsample not selected for the present substudy.

The plasma carotenoids were $\log_e$-transformed, and the correlations between these plasma carotenoids and the identified dietary patterns were estimated by bivariate (Pearson) and partial correlation coefficients (r_{adj}), adjusting for age (continuous), total energy intake (continuous), BMI (continuous), smoking status (yes/no), intake of vitamin supplements (yes/no) and alcohol intake (continuous). The partial correlation coefficients are presented with 95% CI based on 1000 bootstrap samples. The associations between $\log_e$-transformed plasma carotenoids (response variables) and the dietary patterns were further studied by multiple linear regressions. The women's scores for each dietary pattern were categorised into quartiles. The regression models included the same covariates as in the calculation of partial correlation coefficients. To test for trend, the dietary pattern was modelled as a continuous variable in the regression analyses. We report adjusted back-transformed marginal means of plasma carotenoids (μmol/l) with 95% CI.

All tests were two-sided and $P < 0.05$ was considered statistically significant. Statistical analyses were conducted using SPSS version 20.0 (IBM Corporation).

Results

Table 2 presents the selected characteristics of women in the study sample (*n* 361) and the NBCSP subsample (without the study sample, *n* 6120). There was a significant difference in age between the two samples ($P = 0.001$), although the mean values were quite similar (57·1 years in the study sample and 57·9 years in the NBCSP subsample). No significant differences were found regarding BMI, alcohol intake, energy intake and multivitamin use ($0.53 \leq P \leq 0.87$), while the proportion of smokers was significantly lower in the study sample than in the NBCSP subsample (14 *v.* 20%, $P = 0.005$). Furthermore, women in the study sample were significantly higher educated ($P = 0.01$) than those in the NBCSP subsample. Plasma carotenoid levels for the study sample are also presented in Table 2.

The forty-six food groups that formed the basis for the PCA are presented in Table 1. We identified four major dietary patterns, all with eigenvalues ≥ 2.0. The scree plot showed a clear break point, justifying the retaining of the four dietary patterns. The four retained dietary patterns accounted for 6·1, 6·1, 5·7 and 5·1%, respectively, of the total variance. The four dietary patterns with food groups having factor loadings with absolute values ≥ 0.3 are presented in Table 3. The first dietary pattern was labelled 'Western' because of the high positive loadings for potatoes, cakes and desserts, refined grains, high-fat dairy products, sauce, sweet spreads, margarine, sugar-sweetened beverages, cheese, whole grains and low-fat dairy products. The second dietary pattern was labelled 'Vegetarian' because of high loadings for herbs and spices, vegetables, vegetarian food, salad dressings, soya sauce, rice, legumes, nuts and seeds, tea, fruits and berries, and whole grains. The third pattern was labelled 'Continental' because of high loadings for tomato sauce, pasta, fat-rich potatoes, processed meat, pizza, sweets, red meat and game, mustard, salty snacks, wine, and rice. The fourth pattern was labelled 'High-protein' because of high loadings for fish, meat dishes, eggs, vegetables, red meat and game, processed meat, poultry, and mayonnaise.

Table 4 presents the adjusted partial correlation coefficient between the four dietary pattern scores and the plasma carotenoids, and the adjusted means of plasma carotenoids across the quartiles of dietary pattern scores in order to study potential trends. Pearson's correlation coefficients were also calculated, but are not presented since these were similar to the adjusted partial correlation coefficients. The 'Western' pattern scores were significantly inversely correlated with plasma lutein, zeaxanthin, lycopene and total carotenoids ($-0.25 \leq r_{adj} \leq -0.13$). We found a significant decreasing trend across the quartiles of the Western pattern scores only for plasma lycopene ($P_{trend} < 0.001$). Mean lycopene decreased from 0·263 μmol/l in the lowest score quartile to 0·196 μmol/l in the highest score quartile (a 25% decrease). The 'Vegetarian' pattern scores were significantly positively

Table 3. Factor loadings* for the four dietary patterns found in the principal component analysis (*n* 361)

Western		Vegetarian		Continental		High-protein	
Food group	Factor loading	Food group	Factor loading	Food group	Factor loading	Food group	Factor loading
Potatoes	0·55	Herbs and spices	0·50	Tomato sauce	0·50	Fish	0·63
Cakes and desserts	0·52	Vegetables	0·49	Pasta	0·49	Meat dishes	0·54
Refined grains	0·51	Vegetarian food	0·47	Fat-rich potatoes	0·48	Eggs	0·49
High-fat dairy products	0·47	Salad dressings	0·45	Processed meat	0·48	Vegetables	0·45
Sauce	0·43	Soya sauce	0·45	Pizza	0·46	Red meat and game	0·41
Sweet spreads	0·42	Rice	0·42	Sweets	0·41	Processed meat	0·33
Margarine	0·42	Legumes	0·41	Red meat and game	0·34	Poultry	0·33
Sugar-sweetened beverages	0·40	Nuts and seeds	0·41	Mustard	0·33	Mayonnaise	0·32
Cheese	0·36	Tea	0·37	Salty snacks	0·33		
Whole grains	0·32	Fruits and berries	0·34	Wine	0·33		
Low-fat dairy products	0·30	Whole grains	0·34	Rice	0·31		
Wine	−0·36	Coffee	−0·32	Fruits and berries	−0·33		
Poultry	−0·33						
Total variance explained (%)	6·1		6·1		5·7		5·1

* Factor loadings between −0·30 and 0·30 are not listed.

Table 4. Partial correlation coefficients (r_{adj}) between the dietary patterns and plasma carotenoids*, and adjusted mean† levels of plasma carotenoids (μmol/l) across the quartiles (Q) of dietary pattern scores (n 361)‡

(Partial correlation coefficients and 95 % confidence intervals§, and adjusted mean values and 95 % confidence intervals)

	Western		Vegetarian		Continental		High-protein	
	Adjusted mean	95 % CI	Adjusted mean	95 % CI	Adjusted mean	95 % CI	Adjusted mean	95 % CI
Lutein								
r_{adj}	−0·13		0·16		−0·13		0·04	
95 % CI	−0·24, −0·02		0·04, 0·28		−0·26, −0·01		−0·08, 0·15	
Q1	0·204	0·174, 0·238	0·185	0·162, 0·212	0·205	0·181, 0·232	0·197	0·172, 0·225
Q2	0·198	0·172, 0·228	0·189	0·165, 0·215	0·209	0·183, 0·239	0·194	0·172, 0·220
Q3	0·204	0·180, 0·232	0·194	0·170, 0·222	0·201	0·176, 0·231	0·184	0·160, 0·212
Q4	0·188	0·161, 0·218	0·228	0·197, 0·263	0·179	0·155, 0·206	0·220	0·193, 0·251
P_{trend}	0·37		0·01		0·04		0·56	
Zeaxanthin								
r_{adj}	−0·13		0·15		−0·06		0·07	
95 % CI	−0·25, −0·02		0·02, 0·28		−0·20, 0·06		−0·04, 0·18	
Q1	0·042	0·036, 0·050	0·037	0·032, 0·042	0·041	0·036, 0·046	0·040	0·035, 0·046
Q2	0·041	0·035, 0·047	0·040	0·035, 0·046	0·044	0·038, 0·050	0·041	0·036, 0·047
Q3	0·042	0·037, 0·048	0·042	0·036, 0·048	0·041	0·036, 0·047	0·039	0·034, 0·045
Q4	0·040	0·034, 0·047	0·047	0·040, 0·054	0·039	0·034, 0·045	0·045	0·039, 0·051
P_{trend}	0·33		0·01		0·53		0·19	
β-Cryptoxanthin								
r_{adj}	−0·07		0·17		−0·03		−0·05	
95 % CI	−0·16, 0·04		0·08, 0·25		−0·13, 0·06		−0·15, 0·07	
Q1	0·141	0·107, 0·176	0·129	0·099, 0·158	0·151	0·123, 0·179	0·157	0·127, 0·188
Q2	0·150	0·120, 0·182	0·141	0·113, 0·170	0·168	0·138, 0·197	0·166	0·139, 0·196
Q3	0·176	0·148, 0·206	0·163	0·134, 0·194	0·157	0·127, 0·189	0·153	0·122, 0·185
Q4	0·164	0·131, 0·200	0·201	0·169, 0·235	0·156	0·125, 0·188	0·155	0·126, 0·185
P_{trend}	0·93		0·001		0·75		0·68	
α-Carotene								
r_{adj}	−0·03		0·15		−0·13		0·06	
95 % CI	−0·15, 0·08		0·04, 0·24		−0·24, −0·03		−0·07, 0·19	
Q1	0·090	0·072, 0·111	0·091	0·075, 0·109	0·108	0·091, 0·128	0·093	0·077, 0·112
Q2	0·093	0·077, 0·113	0·092	0·077, 0·110	0·112	0·094, 0·135	0·099	0·084, 0·118
Q3	0·111	0·094, 0·133	0·097	0·081, 0·116	0·094	0·078, 0·114	0·108	0·089, 0·131
Q4	0·118	0·096, 0·145	0·134	0·111, 0·163	0·096	0·079, 0·117	0·110	0·092, 0·131
P_{trend}	0·36		0·001		0·05		0·11	
β-Carotene								
r_{adj}	−0·09		0·20		−0·06		−0·07	
95 % CI	−0·19, 0·02		0·08, 0·31		−0·18, 0·05		−0·20, 0·08	
Q1	0·340	0·278, 0·417	0·320	0·269, 0·383	0·356	0·302, 0·420	0·360	0·302, 0·428
Q2	0·336	0·279, 0·404	0·311	0·262, 0·369	0·379	0·319, 0·451	0·355	0·301, 0·419
Q3	0·375	0·318, 0·443	0·329	0·277, 0·390	0·338	0·282, 0·404	0·366	0·305, 0·440
Q4	0·359	0·295, 0·438	0·470	0·390, 0·566	0·338	0·281, 0·407	0·330	0·278, 0·391
P_{trend}	0·73		<0·001		0·26		0·35	
Lycopene								
r_{adj}	−0·25		0·20		0·10		−0·06	
95 % CI	−0·36, −0·14		0·08, 0·31		−0·01, 0·20		−0·17, 0·05	
Q1	0·263	0·220, 0·314	0·215	0·184, 0·251	0·211	0·183, 0·244	0·239	0·205, 0·278
Q2	0·233	0·198, 0·273	0·213	0·183, 0·247	0·220	0·189, 0·256	0·217	0·188, 0·250
Q3	0·231	0·200, 0·267	0·227	0·195, 0·264	0·243	0·208, 0·284	0·249	0·212, 0·292
Q4	0·196	0·165, 0·233	0·268	0·228, 0·315	0·246	0·210, 0·289	0·215	0·185, 0·250
P_{trend}	<0·001		0·01		0·07		0·24	

Table 4. Continued

	Western		Vegetarian		Continental		High-protein	
	Adjusted mean	95 % CI	Adjusted mean	95 % CI	Adjusted mean	95 % CI	Adjusted mean	95 % CI
Total carotenoids								
r_{adj}	−0·15		0·24		−0·08		−0·02	
95 % CI	−0·27, −0·04		0·10, 0·36		−0·20, 0·03		−0·14, 0·11	
Q1	1·130	0·985, 1·296	1·030	0·914, 1·162	1·142	1·022, 1·275	1·133	1·007, 1·275
Q2	1·094	0·967, 1·239	1·037	0·923, 1·164	1·186	1·055, 1·333	1·117	0·999, 1·250
Q3	1·190	1·064, 1·331	1·106	0·985, 1·242	1·115	0·989, 1·259	1·139	1·005, 1·289
Q4	1·124	0·984, 1·285	1·399	1·235, 1·586	1·096	0·968, 1·241	1·149	1·023, 1·289
P_{trend}	0·29		<0·001		0·22		0·91	

* Log_e-transformed.
† Reported means (μmol/l) were back-transformed from the estimated marginal means.
‡ Analyses were adjusted for age (continuous), total energy (continuous), BMI (continuous), smoking status (yes/no), alcohol (continuous) and multivitamin supplements (yes/no).
§ Presented with 95 %CI based on 1000 bootstrap samples.

correlated with all the plasma carotenoids ($0·15 \le r_{adj} \le 0·24$), and with significant increasing trends across the quartiles ($P_{trend} \le 0·01$). Mean total carotenoids increased from 1·030 μmol/l in the lowest score quartile to 1·399 μmol/l in the highest score quartile (a 36 % increase). The 'Continental' pattern scores were inversely correlated with most of the plasma carotenoids, although only significant for lutein and α-carotene ($r_{adj} = -0·13$). We found a significant decreasing trend for both lutein and α-carotene across the quartiles of the 'Continental' pattern ($P_{trend} \le 0·05$). This pattern had a non-significant positive correlation with plasma lycopene ($r_{adj} = 0·10$, $P_{trend} = 0·07$). Finally, no significant correlations between the 'High-protein' pattern and the plasma carotenoids were found ($-0·07 \le r_{adj} \le 0·07$).

Discussion

In the present subsample of middle-aged postmenopausal Norwegian women, we identified four main dietary patterns: 'Western'; 'Vegetarian'; 'Continental'; 'High-protein'. Furthermore, we found that the healthy 'Vegetarian' pattern was associated with a more favourable profile of the plasma carotenoids than the more unhealthy 'Western' and 'Continental' patterns. We found no significant associations between the 'High-protein' pattern and the plasma carotenoids.

Studies on dietary patterns derived from PCA for women in this age group are limited. Dietary patterns will never be exactly the same across studies; however, in most populations, a healthy ('Prudent') pattern and a less healthy ('Western') pattern have been identified[31]. Our 'Western' pattern is, to some extent, similar to 'Western' patterns found in previous studies, which is usually characterised by high loadings of red meat, processed meat, butter, potatoes, refined grains and high-fat dairy products[32–34]. We labelled our 'Prudent' pattern the 'Vegetarian' pattern since all the food groups with high loadings represented in this pattern are plant-based. Similar patterns have been reported by other researchers. A study conducted on active and retired female teachers from California found five dietary patterns, and one, labelled 'Plant-based', was characterised by high loadings of fruit and vegetables[35]. An Australian study of young, middle-aged and older women identified six dietary patterns, of which two, 'Cooked vegetables' and 'Fruit', had high loadings of vegetables and fruits, respectively[36]. Our 'Continental' pattern is, to some extent, similar to the 'Pasta and meat' pattern found in the Italian European Prospective Investigation into Cancer and Nutrition cohort[37], and to the 'High protein, high-fat' pattern found in California[35]. Furthermore, in a study of middle-aged Norwegian women (age 41–55 years, *n* 35 554) on dietary patterns and lifestyle factors, one of the patterns had high loadings of rice, pasta, red meat and meat products, pizza, chocolate, and salty snacks[5].

The ability to measure habitual food intake accurately is difficult when the assessment relies on self-reported dietary data. We used plasma carotenoids as objective biomarkers to see whether the obtained dietary patterns were meaningful. The negative associations that we found between the

'Western' pattern and the plasma carotenoids are as expected, since neither fruits nor vegetables were represented among the food groups with high loadings in this pattern. In contrast, there was a significant increasing trend in the mean total plasma carotenoid values from the lowest to the highest quartile of 'Vegetarian' pattern scores. Comparisons with other studies are limited, as only a few studies have investigated the association between empirically derived dietary patterns and plasma carotenoids. Our findings are consistent with those of a study conducted among men from the Health Professionals Follow-up Study in the USA[38], showing that the 'Prudent' pattern was positively and the 'Western' pattern was negatively associated with plasma carotenoids. Pearson's correlation coefficients between the 'Prudent' pattern and different plasma carotenoids ranged from 0·03 to 0·26, whereas for the 'Western' pattern ranged from −0·26 to −0·06. These estimates are in the same range as those found in the present study. Our findings are also consistent with those of Talegawkar et al.[13] who studied African American men and women in the southern USA and identified four dietary patterns by cluster analysis ('Fast food', 'Southern', 'Prudent' and 'Juice'), and found that subjects belonging to the more unhealthy patterns had lower concentrations of serum carotenoids than those belonging to the more healthy patterns. A study of men and women aged >65 years in the UK reported higher plasma carotenoids among the participants with a healthy dietary pattern compared with those with more unhealthy patterns[39].

For the 'Continental' pattern, the highest factor loading was for tomato sauce. Lycopene is found in high concentrations in tomatoes, and food processing and cooking improves intestinal absorption[40]. As expected, we found an indication of a positive association between plasma lycopene and this pattern, and an increasing trend in the plasma lycopene of women from the lowest to the highest quartile of the 'Continental' pattern scores. Hendrickson et al.[41] developed empirical prediction models (where foods were weighted according to their relation to biological nutrient levels) for plasma carotenoids, and found that the individual food explaining the largest percentage of variation in plasma lycopene was tomato sauce. Maiani et al.[19] reported that in France, the Republic of Ireland, the UK, The Netherlands and Spain, the major foods contributing to lycopene intake in adults are fresh tomatoes, canned tomatoes, tomato soup, pizza and tomato purée, which tend to have a low content of other carotenoids. As in the 'Western' pattern, fruits and vegetables were not among the food groups with high positive loadings in the 'Continental' pattern. In fact, fruits and berries had a high negative loading (−0·33), and we found significant inverse associations between this pattern and plasma lutein and α-carotene.

Finally, for the 'High-protein' pattern, we did not find any associations with the plasma carotenoids, even though vegetables were represented by high factor loadings. The reason for why there was no association between the plasma carotenoids and this pattern is uncertain, but may be due to over-reporting of vegetables. In a study of 3730 pregnant Finnish women, Arkkola et al.[42] found seven dietary patterns. In the 'Traditional bread' pattern, berry juice was the only carotenoid-rich food group represented by factor loadings >0·3; however, it had no impact on the association between this pattern and the plasma carotenoids (r 0·06).

The present study holds some limitations. Validated FFQ are accepted as a tool in nutritional epidemiological investigations[2], despite known measurement errors, both random and systematic[43,44]. As in all studies using self-reported dietary intake data, measurement errors are inevitable. FFQ require that the respondents remember the amount and frequency of intake of a substantial amount of food items, as a result of which there arise numerous cognitive difficulties[45]. A general under-reporting of food intake, especially among overweight participants, is a well-known limitation of self-reported dietary intake[46,47]. Over-reporting of 'healthy foods' related to social desirability has also been reported[48]. A FFQ may also cover less of the total diet of some individuals than others due to the closed food list. Such measurement errors may distort the composition of the food groups in a dietary pattern, and may have made it more difficult to discern distinctive patterns. Whether this could have biased the associations between the dietary patterns and the plasma carotenoids is more difficult to predict. In general, measurement errors tend to obscure the associations.

Many subjective decisions are made when assessing dietary patterns. This can have a direct impact on the number and type of dietary patterns derived. The first decision is how to group the food items, and then whether to adjust the food group variables for energy before entry into the PCA. Furthermore, decision has to be made regarding the criteria for determining how many components to extract (i.e. the eigenvalue cut-off and the break point in the scree plot) as well as the method of rotating the selected components. Northstone et al.[49] found differences between the dietary patterns obtained using unadjusted and energy-adjusted (by the residual method[50]) food group variables in the PCA; however, these differences did not have any major impact on their further analysis of the associations with nutrient intakes and birth weight. They concluded that 'it is sufficient to make energy adjustment when analysing the effects of the dietary patterns on the outcome of interest'. Smith et al.[51] found no benefit associated with using energy-adjusted gram weights or percentage energy contribution as the input variable strategy, and their recommendation was to use weights of foods to describe underlying dietary patterns in a population. Thus, we did not adjust the food group variables for energy. By using the correlation matrix of the forty-six food group variables in the PCA, and not the covariance matrix, the variables were standardised. The three components retained from the PCA accounted for 23·0% of the total variance. This is consistent with the results of other studies that have used comparable methods where the explained variance was 21–26% for four components[37,52–54]. The percentage of the variance of the food group variables that is explained by the components depends on the number of variables included. We had forty-six food groups as input variables to the PCA. With even more aggregation of the food groups, it is likely that we would

have accounted for a higher proportion of the variance, but the level of detail would be reduced, which was undesirable.

When comparing the characteristics of the subjects in the study sample and the NBCSP subsample, we found significant differences in age, education and smoking status. The age in the study sample was 57·1 *v.* 57·9 years in the NBCSP subsample. Although the difference was significant, the difference was rather small, and although theoretically possible, we think it is unlikely that this slight age difference has introduced selection bias in the present study. The proportion of smokers was lower in the study sample than in the NBCSP subsample (14 *v.* 20 %), and the subjects in the study sample were somewhat higher educated. Smokers are known to make poorer food choices and consume a more energy-dense diet that is inferior in nutritional quality to that of non-smoking adults[55,56]. The present results from a slightly 'healthier' population may, therefore, not be generalisable to all women participating in the NBCSP. However, to what extent these differences in smoking and education resulted in any selection bias that affected our findings is not obvious. With a healthier population, we could speculate that the population may be more homogeneous, which may have impaired our ability of finding any dietary patterns, whether healthy or unhealthy.

In conclusion, we identified four dietary patterns: 'Western'; 'Vegetarian'; 'Continental'; 'High-protein'. All the plasma carotenoids were positively associated with the 'Vegetarian' pattern, while inverse associations were observed between most of the plasma carotenoids and the 'Western' and 'Continental' patterns. There were no associations between the 'High-protein' pattern and the plasma carotenoids. These results show that our healthy dietary pattern, the 'Vegetarian' pattern, is associated with a more favourable profile of the plasma carotenoids than our unhealthy dietary patterns, the 'Western' and 'Continental' patterns. Empirically derived dietary patterns are associated with some uncertainty since they are influenced by the measurement errors in the underlying methods of dietary assessment. Therefore, measurement of nutritional biomarkers can be a useful tool for validating dietary patterns.

Acknowledgements

References

1. Michels KB & Schulze MB (2005) Can dietary patterns help us detect diet–disease associations? *Nutr Res Rev* **18**, 241–248.
2. Newby PK & Tucker KL (2004) Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* **62**, 177–203.
3. Lockheart MS, Steffen LM, Rebnord HM, *et al.* (2007) Dietary patterns, food groups and myocardial infarction: a case–control study. *Br J Nutr* **98**, 380–387.
4. Sommer C, Sletner L, Jenum AK, *et al.* (2013) Ethnic differences in maternal dietary patterns are largely explained by socio-economic score and integration score: a population-based study. *Food Nutr Res* **57** (epublication 8 July 2013).
5. Engeset D, Alsaker E, Ciampi A, *et al.* (2005) Dietary patterns and lifestyle factors in the Norwegian EPIC cohort: the Norwegian Women and Cancer (NOWAC) study. *Eur J Clin Nutr* **59**, 675–684.
6. Hann CS, Rock CL, King I, *et al.* (2001) Validation of the Healthy Eating Index with use of plasma biomarkers in a clinical sample of women. *Am J Clin Nutr* **74**, 479–486.
7. Neuhouser ML, Patterson RE, King IB, *et al.* (2003) Selected nutritional biomarkers predict diet quality. *Public Health Nutr* **6**, 703–709.
8. Newby PK, Muller D & Tucker KL (2004) Associations of empirically derived eating patterns with plasma lipid biomarkers: a comparison of factor and cluster analysis methods. *Am J Clin Nutr* **80**, 759–767.
9. Weinstein SJ, Vogt TM & Gerrior SA (2004) Healthy Eating Index scores are associated with blood nutrient concentrations in the third National Health And Nutrition Examination Survey. *J Am Diet Assoc* **104**, 576–584.
10. Kant AK & Graubard BI (2005) A comparison of three dietary pattern indexes for predicting biomarkers of diet and disease. *J Am Coll Nutr* **24**, 294–303.
11. Panagiotakos DB, Pitsavos C, Skoumas Y, *et al.* (2007) The association between food patterns and the metabolic syndrome using principal components analysis: the ATTICA Study. *J Am Diet Assoc* **107**, 979–987; quiz 997.
12. Nettleton JA, Schulze MB, Jiang R, *et al.* (2008) A priori-defined dietary patterns and markers of cardiovascular disease risk in the Multi-Ethnic Study of Atherosclerosis (MESA). *Am J Clin Nutr* **88**, 185–194.
13. Talegawkar SA, Johnson EJ, Carithers TC, *et al.* (2008) Serum carotenoid and tocopherol concentrations vary by dietary pattern among African Americans. *J Am Diet Assoc* **108**, 2013–2020.
14. McNaughton SA, Mishra GD & Brunner EJ (2009) Food patterns associated with blood lipids are predictive of coronary heart disease: the Whitehall II study. *Br J Nutr* **102**, 619–624.
15. Bogl LH, Pietilainen KH, Rissanen A, *et al.* (2013) Association between habitual dietary intake and lipoprotein subclass

profile in healthy young adults. *Nutr Metab Cardiovasc Dis* **23**, 1071–1078.

16. Lipsky LM, Cheon K, Nansel TR, et al. (2012) Candidate measures of whole plant food intake are related to biomarkers of nutrition and health in the US population (National Health and Nutrition Examination Survey 1999–2002). *Nutr Res* **32**, 251–259.

17. Jenab M, Slimani N, Bictash M, et al. (2009) Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Hum Genet* **125**, 507–525.

18. Hammond BR Jr & Renzi LM (2013) Carotenoids. *Adv Nutr* **4**, 474–476.

19. Maiani G, Caston MJ, Catasta G, et al. (2009) Carotenoids: actual knowledge on food sources, intakes, stability and bio-availability and their protective role in humans. *Mol Nutr Food Res* **53**, Suppl. 2, S194–S218.

20. Alberg A (2002) The influence of cigarette smoking on circulating concentrations of antioxidant micronutrients. *Toxicology* **180**, 121–137.

21. Vioque J, Weinbrenner T, Asensio L, et al. (2007) Plasma concentrations of carotenoids and vitamin C are better correlated with dietary intake in normal weight than overweight and obese elderly subjects. *Br J Nutr* **97**, 977–986.

22. Al-Delaimy WK, Ferrari P, Slimani N, et al. (2005) 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* **59**, 1387–1396.

23. Hofvind SS, Wang H & Thoresen S (2003) The Norwegian Breast Cancer Screening Program: re-attendance related to the women's experiences, intentions and previous screening result. *Cancer Causes Control* **14**, 391–398.

24. Hofvind S, Geller B, Vacek PM, et al. (2007) Using the European guidelines to evaluate the Norwegian Breast Cancer Screening Program. *Eur J Epidemiol* **22**, 447–455.

25. Ellingjord-Dale M, dos-Santos-Silva I, Grotmol T, et al. (2015) Vitamin D intake, month at mammography and mammographic density in Norwegian women aged 50–69. Submitted.

26. Andersen LF, Solvoll K, Johansson LR, et al. (1999) Evaluation of a food frequency questionnaire with weighed records, fatty acids, and α-tocopherol in adipose tissue and serum. *Am J Epidemiol* **150**, 75–87.

27. Carlsen MH, Lillegaard IT, Karlsen A, et al. (2010) Evaluation of energy and dietary intake estimates from a food frequency questionnaire using independent energy expenditure measurement and weighed food records. *Nutr J* **9**, 37.

28. Carlsen MH, Karlsen A, Lillegaard IT, et al. (2011) Relative validity of fruit and vegetable intake estimated from an FFQ, using carotenoid and flavonoid biomarkers and the method of triads. *Br J Nutr* **105**, 1530–1538.

29. Tabachnick BG & Fidell FL (2007) *Using Multivariate Statistics*, 5th ed. New York: Pearson Education Publications.

30. Hutcheson G & Sofroniou N (1999) *The Multivariate Social Scientist*. London: Sage.

31. Ocke MC (2013) Evaluation of methodologies for assessing the overall diet: dietary quality scores and dietary pattern analysis. *Proc Nutr Soc* **72**, 191–199.

32. Chocano-Bedoya PO, O'Reilly EJ, Lucas M, et al. (2013) Prospective study on long-term dietary patterns and incident depression in middle-aged and older women. *Am J Clin Nutr* **98**, 813–820.

33. Weismayer C, Anderson JG & Wolk A (2006) Changes in the stability of dietary patterns in a study of middle-aged Swedish women. *J Nutr* **136**, 1582–1587.

34. Nimptsch K, Malik VS, Fung TT, et al. (2014) Dietary patterns during high school and risk of colorectal adenoma in a cohort of middle-aged women. *Int J Cancer* **134**, 2458–2467.

35. Link LB, Canchola AJ, Bernstein L, et al. (2013) Dietary patterns and breast cancer risk in the California Teachers Study cohort. *Am J Clin Nutr* **98**, 1524–1532.

36. Mishra GD, McNaughton SA, Ball K, et al. (2010) Major dietary patterns of young and middle aged women: results from a prospective Australian cohort study. *Eur J Clin Nutr* **64**, 1125–1133.

37. Pala V, Sieri S, Masala G, et al. (2006) Associations between dietary pattern and lifestyle, anthropometry and other health indicators in the elderly participants of the EPIC-Italy cohort. *Nutr Metab Cardiovasc Dis* **16**, 186–201.

38. Hu FB, Rimm E, Smith-Warner SA, et al. (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* **69**, 243–249.

39. Pryer JA, Cook A & Shetty P (2001) Identification of groups who report similar patterns of diet among a representative national sample of British adults aged 65 years of age or more. *Public Health Nutr* **4**, 787–795.

40. Wang XD (2012) Lycopene metabolism and its biological significance. *Am J Clin Nutr* **96**, 1214S–1222S.

41. Hendrickson SJ, Willett WC, Rosner BA, et al. (2013) Food predictors of plasma carotenoids. *Nutrients* **5**, 4051–4066.

42. Arkkola T, Uusitalo U, Kronberg-Kippila C, et al. (2008) Seven distinct dietary patterns identified among pregnant Finnish women – associations with nutrient intake and sociodemographic factors. *Public Health Nutr* **11**, 176–182.

43. Schatzkin A & Kipnis V (2004) Could exposure assessment problems give us wrong answers to nutrition and cancer questions? *J Natl Cancer Inst* **96**, 1564–1565.

44. Kristal AR, Peters U & Potter JD (2005) Is it time to abandon the food frequency questionnaire? *Cancer Epidemiol Biomarkers Prev* **14**, 2826–2828.

45. Subar AF, Thompson FE, Smith AF, et al. (1995) Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. *J Am Diet Assoc* **95**, 781–788; quiz 789–790.

46. Patterson RE, Kristal AR, Tinker LF, et al. (1999) Measurement characteristics of the Women's Health Initiative food frequency questionnaire. *Ann Epidemiol* **9**, 178–187.

47. Neuhouser ML, Tinker L, Shaw PA, et al. (2008) Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol* **167**, 1247–1259.

48. Hebert JR, Clemow L, Pbert L, et al. (1995) Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *Int J Epidemiol* **24**, 389–398.

49. Northstone K, Ness AR, Emmett PM, et al. (2008) Adjusting for energy intake in dietary pattern investigations using principal components analysis. *Eur J Clin Nutr* **62**, 931–938.

50. Willett WC (2013) *Nutritional Epidemiology*, 3rd ed. New York: Oxford University Press.

51. Smith AD, Emmett PM, Newby PK, et al. (2013) Dietary patterns obtained through principal components analysis: the effect of input variable quantification. *Br J Nutr* **109**, 1881–1891.

52. Northstone K & Emmett PM (2010) Dietary patterns of men in ALSPAC: associations with socio-demographic and lifestyle characteristics, nutrient intake and comparison with women's dietary patterns. *Eur J Clin Nutr* **64**, 978–986.

53. Kesse-Guyot E, Bertrais S, Peneau S, et al. (2009) Dietary patterns and their sociodemographic and behavioural

correlates in French middle-aged adults from the SU.VI.MAX cohort. *Eur J Clin Nutr* **63**, 521–528.

54. Lioret S, McNaughton SA, Crawford D, *et al.* (2012) Parents' dietary patterns are significantly correlated: findings from the Melbourne Infant Feeding Activity and Nutrition Trial Program. *Br J Nutr* **108**, 518–526.

55. Cade JE & Margetts BM (1991) Relationship between diet and smoking – is the diet of smokers different? *J Epidemiol Community Health* **45**, 270–272.

56. Dyer AR, Elliott P, Stamler J, *et al.* (2003) Dietary intake in male and female smokers, ex-smokers, and never smokers: the INTERMAP study. *J Hum Hypertens* **17**, 641–654.