

# Dietary patterns and breast cancer risk: results from three cohort studies in the DIETSCAN project

Citation for published version (APA):

Männistö, S., Dixon, L. B., Balder, H. F., Virtanen, M. J., Krogh, V., Khani, B. R., Berrino, F., van den Brandt, P. A., Hartman, A. M., Pietinen, P., Tan, E. S., Wolk, A., & Goldbohm, R. A. (2005). Dietary patterns and breast cancer risk: results from three cohort studies in the DIETSCAN project. *Cancer Causes & Control*, 16(6), 725-733. <https://doi.org/10.1007/s10552-005-1763-7>

## Document status and date:

Published: 01/01/2005

## DOI:

[10.1007/s10552-005-1763-7](https://doi.org/10.1007/s10552-005-1763-7)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

## Dietary patterns and breast cancer risk: results from three cohort studies in the DIETSCAN project

Satu Männistö<sup>1,\*</sup>, L. Beth Dixon<sup>2,7</sup>, Helena F. Balder<sup>3</sup>, Mikko J. Virtanen<sup>1</sup>, Vittorio Krogh<sup>4</sup>, Bahram Rashid Khani<sup>4</sup>, Franco Berrino<sup>4</sup>, Piet A. van den Brandt<sup>6</sup>, Anne M. Hartman<sup>7</sup>, Pirjo Pietinen<sup>1</sup>, Frans Tan<sup>8</sup>, Alicja Wolk<sup>5</sup> & R. Alexandra Goldbohm<sup>3</sup>

<sup>1</sup>Department of Epidemiology and Health Promotion, National Public Health Institute, Mannerheimintie 166, FIN-00300, Helsinki, Finland; <sup>2</sup>Department of Nutrition, Food studies and Public Health, New York University, New York, NY, USA; <sup>3</sup>Department of Nutritional Epidemiology, TNO Nutrition and Food Research, Zeist, The Netherlands; <sup>4</sup>Epidemiology Unit, Istituto Nazionale per lo Studio e la Cura dei Tumori, Milan, Italy; <sup>5</sup>Institute of Environmental Epidemiology, Karolinska Institutet, Stockholm, Sweden; <sup>6</sup>Department of Epidemiology, Maastricht University, Maastricht, The Netherlands; <sup>7</sup>Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA; <sup>8</sup>Department of Methodology and Statistics, Maastricht University, Maastricht, The Netherlands

Received 15 August 2004; accepted in revised form 4 February 2005

**Key words:** breast cancer, diet, dietary pattern, factor analysis, principal component analysis.

### Abstract

**Objective:** Only a few consistent findings on individual foods or nutrients that influence breast cancer risk have emerged thus far. Since people do not consume individual foods but certain combinations of them, the analysis of dietary patterns may offer an additional aspect for assessing associations between diet and diseases such as breast cancer. It is also important to examine whether the relationships between dietary patterns and breast cancer risk are consistent across populations.

**Methods:** We examined the risk of breast cancer with two dietary patterns, identified as “Vegetables” (VEG) and “Pork, Processed Meat, Potatoes” (PPP), common to all cohorts of the DIETSCAN project. During 7 to 13 years of follow-up, three of the cohorts – the Netherlands Cohort Study on diet and cancer (NLCS), the Swedish Mammography Cohort (SMC), and the Ormoni e Dieta nella Eziologia dei Tumori (Italy-ORDET) – provided data on 3271 breast cancer cases with complete information on their baseline diet measured by a validated food frequency questionnaire.

**Results:** After adjustment for potential confounders, VEG was not associated with the risk of breast cancer across all cohorts. PPP was also not associated with the risk of breast cancer in SMC and ORDET, but a high PPP score tended to be inversely associated with breast cancer in the NLCS study (RR = 0.69; 95% CI, 0.52–0.92, highest versus lowest quartile). PPP differed in one aspect between the cohorts: butter loaded positively on the pattern in all cohorts except NLCS, in which butter loaded negatively and appeared to be substituted by low-fat margarine loading positively.

**Conclusion:** In general, the dietary patterns showed consistent results across the three cohorts except for the possible protective effect of PPP in the NLCS cohort, which could be explained by a difference in that pattern for NLCS. The results supported the suggestion derived from traditional epidemiology that relatively recent diet may not have an important role in the etiology of breast cancer.

**Abbreviations:** ATBC – Alpha-Tocopherol Beta-Carotene Cancer Prevention Study; DIETSCAN – DIETary patternS and CANcer in four European countries; FFQ – Food frequency questionnaire; NLCS – Netherlands Cohort Study on Diet and Cancer; ORDET – Ormoni e Dieta nella Eziologia dei Tumori; SMC – Swedish Mammography Cohort

\*Address correspondence to: Satu Männistö, PhD, Senior Researcher, National Public Health Institute, Department of Epidemiology and Health Promotion, Mannerheimintie 166, FIN-00300 Helsinki, Finland. Fax: +358-9-4744 8591; E-mail: satu.mannisto@ktl.fi

## Introduction

Breast cancer is the most common female malignancy in the world, and in recent decades its incidence rate has consistently increased [1]. Much of the international variation in breast cancer rates is explained by differences in reproductive and hormonal factors. Although it has been estimated that dietary factors may account for about 30% of cancers in industrialized countries [2], studies on diet and breast cancer risk have uncovered few definite effects and left much uncertainty [3]. Alcohol consumption increases the risk of breast cancer [4, 5], but the findings for other dietary variables including meat, dairy products, fruit, vegetables, fat, fiber and phytoestrogens are inconsistent [3].

Most epidemiological studies on diet and cancer are based on intakes of individual food items or nutrients. This type of approach, however, does not take into account the complexity of human diet, such as the large number of candidate foods, and intercorrelation between dietary components [6]. As a consequence, the analysis of dietary patterns offers an additional dimension to assess associations between diet and diseases [7–9]. Overall, dietary patterns may have a greater effect on health than any individual food item or nutrient, and are more useful in determining public health recommendations [10].

Exploratory factor analysis is a statistical method to define dietary patterns. The aim of the method is to compress detailed dietary data into a few specific dietary patterns by analyzing their covariate structure [11]. Thus far, certain dietary patterns have been significantly associated with total mortality [12], coronary heart disease [8, 13], diabetes [9] and different cancer sites [7, 14–20]. Dietary patterns have also been related to biochemical indicators [8, 21], and they are more effective than specific nutrients in clinical trials [22, 23]. Thus, analysis of dietary patterns is a good candidate approach for further research in order to complete the findings of traditional nutritional epidemiology [10].

Exploratory factor analysis has been criticized because the dietary patterns may be unique to a single population but not to be applicable across other populations [24, 25]. To date, no study has examined whether dietary patterns and risk of breast cancer are consistent across different populations. The primary aim of the greater DIETSCAN project was to develop and apply a common methodological approach to study dietary patterns and cancer in four European cohort studies: the Alpha-Tocopherol Beta-Carotene Cancer Prevention Study in Finland (ATBC), the Netherlands Cohort Study on diet and cancer (NLCS), the Ormoni e Dieta

nella Eziologia dei Tumori in Italy (ORDET), and the Swedish Mammography Cohort (SMC) [26]. The specific study described in this paper examined whether dietary patterns derived from a common approach and risk of breast cancer are consistent across different populations; and whether dietary patterns contribute additional information to the investigation of diet and breast cancer over looking at single nutrients or foods reported previously in the literature. This study on breast cancer risk included the three of cohorts (NLCS, ORDET and SMC) with female participants, and thus provided breast cancer data. The data were not pooled but results were analyzed in each cohort separately using the same predefined strategy.

## Subjects and methods

### *Study population*

Four European cohort studies included in the greater DIETSCAN project were selected because they met the following criteria: prospective cohort study designed to investigate the effects of diet on risk of various cancers, complete dietary assessment including the entire diet, and a validation study of the dietary assessment method. Three of these cohort studies (NLCS, ORDET and SMC) who had female participants were included in this specific breast cancer study (Table 1). The NLCS study on diet and cancer is a population-based cohort study of 62,573 women (and 58,279 men), selected from 204 Dutch municipalities [27]. The participants were between 55 and 69 years of age when the study began in 1986. Because of the case-cohort design, a random subcohort of 1812 women (and 1688 men) was sampled from the cohort after the baseline exposure measurement. Because of missing or inconsistent dietary data, the analyses are based on data of 3123 subcohort participants (1598 women and 1525 men). Only the 1598 women in the NLCS are used in this analysis of diet patterns and breast cancer risk. The ORDET study is a cohort study on hormonal factors and diet related to breast cancer risk in Italian women, aged 34 to 70 years [28]. The cohort consisted of 10,788 healthy volunteers recruited between 1987 and 1992. The SMC study is comprised of 66,651 women from two counties in central Sweden [16, 29]. The participants were between 40 and 74 years of age when they were invited to participate in a population-based mammography screening program in 1987 to 1990.

All invasive breast cancer cases were identified through national or local cancer registers and were

Table 1. Characteristics of three DIETSCAN cohort studies with invasive breast cancer cases

	NLCS <sup>a</sup>	ORDET <sup>b</sup>	SMC <sup>c, d</sup>
Country	The Netherlands	Italy	Sweden
Baseline years	1986	1987–1992	1987–1990
Follow-up years	7	9	13
Baseline cohort size	62,573 <sup>e</sup>	10,788	61,463
Age range	55–69	35–69	40–74
Body mass index (kg/m <sup>2</sup> )	25.1 + 3.5	25.4 + 4.3	24.8 + 4.4
High education <sup>f</sup> (yes, %)	31	28	20
Age at menarche (years)	13.7 + 1.8	12.9 + 1.5	13.4 + 1.4
Age at menopause (years)	49 + 4	48 + 5	49 + 5
Parity (yes, %)	83	89	85
Age at first birth (years)	22 + 11	26 + 5	24 + 5
Ever use of oral contraceptive (yes, %)	25	33	46 <sup>d</sup>
Ever use of hormone replacement treatment (yes, %)	15	7	19 <sup>d</sup>
Family history of breast cancer (yes, %)	8	7	7
History of benign breast diseases (yes, %)	9	33 <sup>g</sup>	–
# Breast cancer cases among all women	1127	212	1932
# Person-years	11,264 <sup>e</sup>	82,114	741,458
# Breast cancer cases among post-menopausal women	1127	96	–
# Person-years	11,264 <sup>e</sup>	30,480	–

<sup>a</sup> NLCS – Netherlands Cohort Study on Diet and Cancer (The Netherlands).

<sup>b</sup> ORDET – Ormoni e Dieta nella Etiologia dei Tumori (Italy).

<sup>c</sup> SMC – Swedish Mammography Cohort (Sweden).

<sup>d</sup> Information was based on a sub-cohort of 26,062 women.

<sup>e</sup> A random subcohort of 3123 participants (1598 women) was sampled from the cohort after the baseline exposure measurement.

<sup>f</sup> High school, vocational school or university.

<sup>g</sup> The cohorts consisted of healthy volunteers.

histologically confirmed. In all, the cohorts included 3271 breast cancer cases with complete information on their diet (Table 1).

The NLCS study was approved by the institutional review boards of the TNO Toxicology and Nutrition Institute (Zeist) and the University of Maastricht in the Netherlands. The ORDET was approved by the Ethical Review Board of the Italian National Cancer Institute in Italy. The SMC study was approved by the Institutional Ethics Committee at Uppsala University and at Karolinska Institute (Stockholm) in Sweden.

#### *Food grouping and assessment of dietary patterns*

The validated semi-quantitative FFQs used to collect dietary data at baseline [28, 30, 31] differed in the number of food items, reference period, units of frequency and quantification of portion size, and overall level of detail (Table 2). To achieve the objective of the DIETSCAN project, a common food grouping was developed [26]. Food items from the FFQs were aggregated into 51 food groups that included foods common to all countries as well as specific foods included in each FFQ. These food groups were selected because of their role in the diet and possible relevance to cancer etiology.

To identify dietary patterns within each cohort separately, exploratory factor analysis using principal component analysis was conducted using the average daily amount (gram) consumed of the FFQ-derived food groups for each country. Factor analysis aggregates correlated variables as uncorrelated factors (patterns), explaining as much of the variation in the original variables as possible. Those food groups that were extremely skewed (more than 75% non-users) were dichotomized. Factors were rotated by orthogonal Varimax transformation. The Scree test, the traditional criteria in factor analysis, was used to determine the number of factors to extract within each cohort (*i.e.*, 5 for NLCS, 4 for ORDET and 4 for SMC) [26].

While labeling the factors, food groups with absolute factor loadings  $\geq 0.35$  were considered as contributing substantially to a dietary pattern (Appendix Table A). Factor loadings represent correlation coefficients between food groups and dietary patterns. Food groups with positive loadings are positively associated with a dietary pattern; food groups with negative loadings are inversely associated with a dietary pattern. Names of dietary patterns (for illustration, “Vegetables” and “Pork, processed meat, potatoes” shown in Appendix Table A) within each cohort were determined according

Table 2. Characteristics of the FFQs and factor analysis-derived dietary patterns in the DIETSCAN cohort studies

	NLCS	ORDET	SMC
Total number of items	150	107	67
FFQ reference period	12 months	12 months	6 months
Frequency	7 categories (never – 6–7/wk)	Times per week/month	8 categories (never/seldom – 4 + /d)
Quantification	Natural or household units (fixed weight per unit)	Portion size pictures (less/equal/more than 1–3 pictures)	Age-specific standard portion sizes
Energy intake (MJ/day)	7.1 + 1.7	7.4 + 2.2	5.6 + 1.6
# Food groups used to determine factors <sup>a</sup>	49	31	42
# Factors extracted	5	4	4
Factor (variance)			
Vegetables	Vegetables (6.3%)	Vegetables (11.6%)	Vegetables (6.9%)
Pork, Processed meat, Potatoes (PPP)	PPP (4.3%)	PPP (7.4%)	PPP (5.4%)
Other factors	Fat dairy (4.4%) Brown/white bread substitute (4.3%) Sweet and savoury snacks (3.9%)	Other cooked vegetables (5.3%) Alcohol (4.7%)	Alcohol (5.3%) Margarine/butter substitution (4.2%)

<sup>a</sup> The same food grouping was used across the cohort studies (food items from the FFQs were aggregated into the 51 food groups). However, because of some missing food items, the true number of aggregated food groups was below 51 in each cohort study.

to the dominant foods (*i.e.*, foods with high loadings) in the respective patterns. Further details regarding the food grouping and dietary pattern assessment, and related sensitivity analyses conducted in the DIETSCAN study, are provided by Balder *et al.* [26].

### Risk analyses

Within ORDET and SMC, Cox proportional hazard models were constructed to estimate hazard ratios and 95% confidence intervals (95% CI) relating the common factors to the incidence of breast cancer. NLCS used survival analysis with exponential distribution to estimate the standard errors using the robust option to account for additional variance introduced by the case-cohort design. Prior to statistical analysis, factor scores were determined using the regression method by Bartlett [32]. Thus, within each cohort, each participant had a unique score for each factor. High scores represented high intake of food items with positive loading on the corresponding dietary pattern while low scores represented low intake of those items. The factor scores were entered into the statistical models as continuous (*i.e.*, one estimate per factor), as grouped (*i.e.*, quartiles), and grouped continuous with medians (*i.e.*, each group assigned the median value within each group) to test for linear trend.

The models were adjusted for other factors, including age, energy intake, and other potential confounding variables (listed in Tables 3 and 4) determined by previous analyses of diet and breast cancer risk conducted within each cohort [16, 28, 29, 33]. Previous analysis [26]

showed that energy adjustment of the food group variables by the residual method yielded comparable factor solutions, so the study-specific factor scores remained unadjusted but energy intake was added to the models as a continuous variable. All reported p values are two-sided.

### Results

The participants of NLCS were on average older ( $61.4 \pm 4.3$  years) than the women in the two other cohorts ( $48 \pm 8.5$  years in ORDET and  $53.7 \pm 9.7$  in SMC) (Table 1). The participants of SMC had more lifetime oral contraceptive use, whereas the participants of ORDET had less lifetime hormone replacement therapy, and had later age at first birth than participants in the other cohorts. A third of ORDET participants had a history of benign breast disease; they were volunteers recruited from the general population through public meetings, by advertising, and among women attending breast cancer prevention units for early diagnosis.

Using the same factor analysis procedure, we identified four to five dietary patterns in the NLCS (5), ORDET (4) and SMC (4) explaining 23.2%, 29.0% and 21.8% of the total variance in the original dietary variables, respectively (Table 2). Two of the dietary patterns were consistent across the cohorts. The first (Vegetables – VEG) pattern was characterized by high intakes of vegetables, legumes, fruit, pasta, fish and oil (Appendix Table A). VEG contributed to the total variance, ranging from 6.3%

Table 3. Characteristics of the vegetable pattern and the pork, processed and potatoes pattern

Cohort study	Vegetables pattern (VEG)		Pork, processed meat and potatoes pattern (PPP)	
	Quartile 1, mean (g/day)	Quartile 4, mean (g/day)	Quartile 1, mean (g/day)	Quartile 4, mean (g/day)
<i>NLCS</i>				
Total fat	66	82	72	80
Saturated fatty acids	27	33	32	30
Fiber	21	31	25	27
Ethanol	3.3	8.2	6.2	5.2
Fruits	161	235	223	180
Vegetables	121	293	194	212
<i>ORDET</i>				
Total fat	54	84	50	90
Saturated fatty acids	20	28	16	33
Fiber	16	24	17	23
Ethanol	8.7	10	5.5	14
Fruits	113	313	199	220
Vegetables	240	396	351	298
<i>SMC</i>				
Total fat	46	46	39	54
Saturated fatty acids	20	19	16	23
Fiber	13	22	17	18
Ethanol	2.6	3.3	3.0	3.1
Fruits	72	259	165	153
Vegetables	41	151	79	101

in NLCS to 11.6% in ORDET (Table 2). In the ORDET study, another unique factor was extracted covering mainly “other” cooked vegetables. The second (Pork, Processed meat, Potatoes – PPP) pattern was characterized by higher intakes of pork, beef, processed meats, potatoes, rice, poultry, liver, butter/low-fat margarine, pasta and coffee (Appendix table A). PPP contributed from 4.3% in NLCS to 7.4% in ORDET (Table 2) to the total variance. Table 3 shows the differences in food and nutrient intake for each factor. The total amounts consumed of these foods and nutrients (g/day) are usually associated with daily energy intake (Table 2).

There were no significant associations between VEG and breast cancer risk in any of the cohorts when all breast cancer cases were included (Table 4). The multivariate continuous models and quartile models of VEG showed very similar risks compared with the models adjusted only for age and energy; RR = 0.90 (95% CI, 0.67–1.20, comparison of the highest versus lowest quartile) in NLCS, RR = 0.79 (95% CI, 0.50–1.27) in ORDET, and RR = 0.93 (95% CI, 0.81–1.07) in SMC. When postmenopausal women were analyzed separately in ORDET, the results for VEG were similar to the overall results based on all women (data not shown).

In the age- and energy-adjusted model, PPP was not associated with the risk of breast cancer in ORDET and SMC (Table 5). Further adjustment for other factors related to breast cancer did not change these results; RR = 1.07 (95% CI, 0.58–1.98, comparison of the highest versus lowest quartile) in ORDET, and RR = 0.92 (95% CI, 0.78–1.09) in SMC. However, an inverse association between PPP and breast cancer risk was found in NLCS. When PPP was used as a continuous variable, the risk of the multivariate model was 0.90 per unit of the factor score (95% CI, 0.81–0.99), whereas the risk of breast cancer was 0.69 (0.52–0.92) when the highest quartile of PPP was compared with the lowest quartile. When postmenopausal women were analyzed separately in ORDET, PPP was not associated with the risk of breast cancer (data not shown).

### Discussion

Two common dietary patterns, one consisting mostly of vegetables (VEG), and one of pork, processed meat and potatoes, were identified with the same factor analysis procedure across three cancer cohorts in the Netherlands, Italy and Sweden [26]. Similar to previous studies [7, 18], VEG represents individuals with a healthier diet and lifestyle, whereas individuals with a high PPP pattern score have more traditional or less healthy lifestyle factors. In general, the observed dietary patterns were not associated with the risk of breast cancer, although a high PPP pattern score was protective for breast cancer in the NLCS cohort.

High consumption of fruit and vegetables rich in fiber, antioxidants, flavonoids and phytoestrogens has been related to decreased risk of all cancers [34, 35]. The association, however, has been more consistent for cancers of the lung and colon than for hormone-related cancers. The previous studies from NLCS and SMC did not find a strong role for the intake of vegetables, fruit, fiber or vitamins in the etiology of breast cancer [29, 36, 37]. A pooled analysis of eight large cohort studies also did not strongly associate high intake of fruit and vegetables with the risk of breast cancer [38]. Our study supported the results that a dietary pattern rich in vegetables and fruits but also characterized by other associated foods as oil and fish, generally considered healthy, was not significantly related to decreased risk of breast cancer.

Epidemiological studies have been inconsistent in their findings concerning breast cancer risk and specific high-fat foods, such as meat [39–41]. No association was observed between meat consumption and breast cancer risk in a pooled analysis of eight cohort studies [42]. Because of the findings that high consumption of

Table 4. Relative risks of breast cancer by linear and quartile models of the vegetables (VEG) pattern in the DIETSCAN cohort studies

	Continuous RR <sup>a</sup>	95% CI	Q1 RR	Q2 RR	95% CI	Q3 RR	95% CI	Q4 RR	95% CI	<i>p</i> Value for trend
<i>NLCS</i>										
Cases	1127		279	292		277		279		
Person-years <sup>b</sup>	11,264		2821	2748		2825		2869		
Age- and energy-adjusted model	0.95	0.87–1.04	1.00	1.07	0.86–1.33	0.97	0.77–1.21	0.94	0.74–1.19	0.44
Multivariate model <sup>c</sup>	0.93	0.84–1.04	1.00	1.21	0.93–1.57	1.03	0.78–1.35	0.90	0.67–1.20	0.31
<i>ORDET</i>										
Cases	210		62	55		51		42		
Person-years	82,114		20,522	20,520		20,618		20,454		
Age- and energy-adjusted model	0.88	0.73–1.07	1.00	0.94	0.65–1.36	0.89	0.60–1.32	0.78	0.49–1.24	0.28
Multivariate model <sup>d</sup>	0.88	0.72–1.07	1.00	0.95	0.65–1.38	0.89	0.60–1.33	0.79	0.50–1.27	0.32
<i>SMC</i>										
Cases	1932		499	463		484		480		
Person-years	741,458		188,177	186,712		184,628		181,939		
Age- and energy-adjusted model	0.99	0.93–1.04	1.00	0.91	0.80–1.04	0.95	0.83–1.08	0.95	0.83–1.09	0.61
Multivariate model <sup>e</sup>	0.97	0.91–1.03	1.00	0.89	0.79–1.02	0.93	0.82–1.06	0.91	0.79–1.05	0.19

<sup>a</sup> The factor scores in the linear models are standardized variables with a mean of zero and a standard deviation of one.

<sup>b</sup> Person-years in NLCS are from the subcohort only.

<sup>c</sup> Multivariate model includes age, body mass index, height, education, smoking (current smoking status, number of cigarettes/day, years smoked), family history of breast cancer, age at menarche, age at menopause, age at first birth, ever use of oral contraceptive, ever use of hormone replacement therapy, alcohol intake and energy.

<sup>d</sup> Multivariate model includes age, body mass index, height, education, smoking (status), family history of breast cancer, ever use of oral contraceptive, ever use of hormone replacement therapy, alcohol intake and energy.

<sup>e</sup> Multivariate model includes age, BMI, education, family history of breast cancer, age at first birth, parity, alcohol intake and energy.

Table 5. Relative risks of breast cancer by linear and quartile models of the PPP pattern in the DIETSCAN cohort studies

	Continuous RR <sup>a</sup>	95% CI	Q1 RR	Q2 RR	95% CI	Q3 RR	95% CI	Q4 RR	95% CI	<i>p</i> Value for trend
<i>NLCS</i>										
Cases	1127		296	293		288		250		
Person-years <sup>b</sup>	11,264		2709	2753		2769		3033		
Age- and energy-adjusted model	<b>0.89</b>	<b>0.83–0.96</b>	1.00	0.96	0.78–1.19	0.94	0.75–1.16	<b>0.72</b>	<b>0.58–0.91</b>	<b>0.01</b>
Multivariate model <sup>c</sup>	<b>0.90</b>	<b>0.81–0.99</b>	1.00	0.91	0.70–1.18	0.97	0.74–1.27	<b>0.69</b>	<b>0.52–0.92</b>	<b>0.02</b>
<i>ORDET</i>										
Cases	210		55	61		50		44		
Person-years	82,114		20,371	20,351		20,648		20,744		
Age- and energy-adjusted model	0.92	0.70–1.21	1.00	1.24	0.83–1.85	1.08	0.67–1.72	1.06	0.57–1.96	0.98
Multivariate model <sup>d</sup>	0.93	0.70–1.22	1.00	1.25	0.83–1.86	1.08	0.67–1.74	1.07	0.58–1.98	0.95
<i>SMC</i>										
Cases	1932		517	471		496		448		
Person-years	741,458		186,604	181,034		187,648		186,170		
Age- and energy-adjusted model	1.01	0.94–1.10	1.00	0.95	0.84–1.08	1.00	0.88–1.15	0.91	0.77–1.08	0.71
Multivariate model <sup>e</sup>	1.02	0.94–1.14	1.00	0.95	0.84–1.08	1.02	0.88–1.15	0.92	0.78–1.09	0.47

<sup>a</sup> The factor scores in the linear models are standardized variables with a mean of zero and a standard deviation of one.

<sup>b</sup> Person-years in NLCS are from the subcohort only.

<sup>c</sup> Multivariate model includes age, body mass index, height, education, smoking (current smoking status, number of cigarettes/day, years smoked), family history of breast cancer, age at menarche, age at menopause, age at first birth, ever use of oral contraceptive, ever use of hormone replacement therapy, alcohol intake and energy.

<sup>d</sup> Multivariate model includes age, body mass index, height, education, smoking (status), family history of breast cancer, ever use of oral contraceptive, ever use of hormone replacement therapy, alcohol intake and energy.

<sup>e</sup> Multivariate model includes age, BMI, education, family history of breast cancer, age at first birth, parity, alcohol intake and energy.

olive oil has been associated with decreased risk of breast cancer [43, 44], it has been suggested that the type of fat may be more relevant in the development of breast cancer than the total amount of fat. The SMC study found an inverse association between monounsaturated fatty acids and breast cancer [31], whereas total fat and animal protein were positively associated with the risk of postmenopausal breast cancer in ORDET [28]. In the DIETSCAN project, the unexpected result of PPP as a protective factor against breast cancer in NLCS may be explained by the type of fat consumed; butter that is rich in saturated fatty acids loaded negatively and appeared to be substituted by low-fat margarine, which is rich in unsaturated fatty acids and consequently healthier.

Although factor analysis has its strengths, it has also been criticized for the lack of generally agreed criteria [45]. The method is based on many subjective decisions on the selection of food groups including the number of factors extracted, the method of rotation and the labeling of dietary patterns; all are factors that can affect the interpretation of results [24]. Although we used a standardized procedure across all cohorts and used the same food grouping [26], there were some differences between the analyses concerning, for example, some missing foods in FFQs. In the Western New York Diet Study, the number of food groups did not affect the number or character of the dietary patterns. The reduced information on diet, however, attenuated the association between the healthy dietary pattern and endometrial cancer risk but not the association for the high-fat pattern [14]. The ORDET study, however, found a stronger inverse association (RR = 0.66; 95% CI, 0.46–0.94) between risk of breast cancer and the salad vegetables pattern (mainly consisting of mixed vegetables in salad, raw leafy vegetables, raw tomatoes, raw carrots and olive oil) based on a larger number of food groups than used in the DIETSCAN project [19]. The longer list of food groups may better identify the cultural dietary habits than the shorter list, for example, by separating raw from cooked tomatoes and other vegetable oils from olive oil (cooked tomatoes and olive oil are foods used in traditional Italian dishes). In our study, however, a major task was to ensure that the endpoints, confounders and dietary data were as similar as possible across all the cohort studies [26].

The identified common dietary patterns seemed to agree with the patterns found in previous studies [11]. The common patterns were reproducible in all four cohorts (including the ATBC cohort which did not contribute breast cancer cases to this study) [26, 46], and were identified in one-half of each data set and reproduced in the other half of the respective data set [26].

Although the simple age- and energy-adjusted models were uniform across the cohorts, the multivariate models reflected differences in sociodemographic and behavioral data available from each cohort. For example, SMC did not collect data on pre/postmenopausal status, physical activity, and smoking. It is also important to keep in mind that patterns common to different cohorts include some differences in the consumption of specific foods.

All cohorts in this study used a validated semi-quantitative FFQ to assess the participants' whole diet over the previous 6 (SMC) or 12 months (NLCS and ORDET). Although the FFQ has proved to be a reasonably good method in large epidemiological studies, the measured dietary variables may involve substantial error [47–49]. It is possible that the dietary pattern approach attenuates the associations between dietary factors and diseases compared to the approach with individual food items. This could happen if one or more specific foods or nutrients are related to cancer risk. However, if it is the overall combination of foods (*i.e.*, dietary patterns) which is related to risk, the pattern approach would show the strongest relationships. Two studies have found reasonable reliability and validity of the major dietary patterns defined by factor analysis with data from FFQs [50, 51]. The authors concluded that the dietary pattern approach might be used in nutritional epidemiology as an alternative method of dietary assessment.

Another challenge in investigating breast cancer is to define the most sensitive period in a woman's lifetime in terms of development of disease. It has been speculated that the period between menarche and the first full-term pregnancy is particularly important because of the changes in the breast tissue during that time [52]. Food frequency questionnaires used in large cohort studies have not typically covered that period of life.

Statistical analyses have a key role in nutritional epidemiology to support the objectivity in the interpretation of results [24]. Less common methods, such as factor analysis, are useful to combine the information on many dietary variables to describe associations of overall dietary patterns with diseases [11]. The DIETSCAN project, however, did not associate VEG or generally PPP with the risk of breast cancer. The results supported the suggestion derived from traditional epidemiology that adult diet may not have an important role in the etiology of breast cancer.

#### Acknowledgements

DIETSCAN was supported by a grant from the Commission of the European Communities (QLK1-

CT-1999-00575). NLCS was supported in part by a grant from the Dutch Cancer Society (TNO V 1999-1963). ORDET was supported by the Italian League against Cancer (Milan section) and the Italian Ministry of Health, with contributions from the Italian National Research Council and the Italian Association for Cancer Research. SMC was supported by grants from the Swedish Cancer Society, The World Cancer Research Fund, and the Swedish Foundation for International Cooperation in Research and Higher Education. SM was supported by grants from Juho Vainio Foundation.

## Appendix

Table A. Factor loads derived from food frequency questionnaires for food groups that represent the common patterns in the DIET-SCAN project<sup>a</sup>

Food groups	NLCS Load	ORDET Load	SMC Load
<i>Vegetables</i>			
Legumes	<b>0.55</b>	0.19	-0.01
Cabbages	<b>0.59</b>	0.22	<b>0.43</b>
Leaf vegetables – cooked	<b>0.48</b>	0.29	0.25
Leaf vegetables – raw	<b>0.52</b>	<b>0.78</b>	<b>0.63</b>
Allium vegetables	<b>0.59</b>	<sup>b</sup>	<sup>b</sup>
Carrots	<b>0.50</b>	<b>0.42</b>	<b>0.59</b>
Tomatoes	<b>0.45</b>	<b>0.72</b>	<b>0.56</b>
Mushrooms	<b>0.39</b>	<sup>b</sup>	<sup>b</sup>
Citrus fruit	0.18	0.32	<b>0.50</b>
Apples and pears	0.17	<sup>b</sup>	<b>0.51</b>
Rice	<b>0.35</b>	-0.02	0.18
Pasta	<b>0.36</b>	0.06	0.02
Fish	0.33	0.11	<b>0.41</b>
Oil	<b>0.41</b>	<b>0.63</b>	<sup>b</sup>
Dressings	0.16	<b>0.72</b>	<sup>b</sup>
<i>Pork, processed meat, potatoes</i>			
Pork	<b>0.43</b>	0.31	<b>0.57</b>
Processed meat	<b>0.50</b>	<b>0.39</b>	<b>0.42</b>
Potatoes	0.29	<b>0.51</b>	0.15
Coffee	<b>0.44</b>	0.11	0.08
Beef and veal	-0.07	<b>0.49</b>	<b>0.60</b>
Pasta	-0.05	<b>0.55</b>	<b>0.55</b>
Rice	-0.31	0.31	<b>0.47</b>
Poultry	0.02	0.04	<b>0.40</b>
Liver	0.28	0.18	<b>0.40</b>
Butter	<b>-0.45</b>	<b>0.61</b>	0.02
Low-fat margarine	<b>0.50</b>	<sup>b</sup>	<sup>b</sup>

<sup>a</sup> Food groups that are listed for each pattern have factor loads  $\geq |0.35|$  in at least one of the DIETSCAN cohort studies with female participants.

<sup>b</sup> Food groups missing from the FFQ and thus, not entered into the principal component analysis.

## References

- Key TJ, Verkasalo PK, Banks E (2001) Epidemiology of breast cancer. *Lancet Oncol* **2**: 133–140.
- WHO. Diet, nutrition and the prevention of chronic diseases (2003). WHO Technical Report Series 916.
- Key TJ, Allen NE, Spencer EA, Travis RC (2002) The effects of diet on risk of cancer. *Lancet* **360**: 861–868.
- Smith-Warner SA, Spiegelman D, Yaun SS, *et al.* (1998) Alcohol and breast cancer in women: a pooled analysis of cohort studies. *J Am Med Assoc* **279**: 535–540.
- Collaborative Group on Hormonal Factors in Breast Cancer (2002) Alcohol, tobacco and breast cancer - collaborative reanalysis of individual data from 53 epidemiological studies, including 58515 women with breast cancer and 95067 women without the disease. *Br J Cancer* **87**: 1234–1245.
- Willett W. (1998) *Nutritional Epidemiology*, 2nd edn. Oxford University Press., Oxford .
- Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma Khe-Ni (1998) Eating patterns and risk of colon cancer. *Am J Epidemiol* **148**: 4–16.
- Fung TT, Rimm EB, Spiegelman D, *et al.* (2001) Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Clin Nutr* **73**: 61–67.
- Dam RMvan, Rimm EB, Willett WC, Stampfer MJ, Hu FB (2002) Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. *Ann Intern Med* **136**: 201–209.
- Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* **13**: 3–9.
- Schulze MB, Hu FB (2002) Dietary patterns and risk of hypertension, type 2 diabetes mellitus, and coronary heart disease. *Curr Atheros Rep* **4**: 462–467.
- Osler M, Heitmann BL, Gerdes LU, Jorgensen LM, Schroll M (2001) Dietary patterns and mortality in Danish men and women: a prospective observational study. *Br J Nutr* **85**: 219–225.
- Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D, Willett WC (2000) Prospective study of major dietary patterns and risk of coronary heart disease in men. *Am J Clin Nutr* **72**: 912–921.
- 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 Nutr* **4**: 989–997.
- Terry P, Hu FB, Hansen H, Wolk A (2001) Prospective study of major dietary patterns and colorectal cancer risk in women. *Am J Epidemiol* **154**: 1143–1149.
- Terry P, Suzuki R, Hu FB, Wolk A (2001) A prospective study of major dietary patterns and the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* **10**: 1281–1285.
- Handa K, Kreiger N (2002) Diet patterns and the risk of renal cell carcinoma. *Public Health Nutr* **5**: 757–767.
- Fung T, Hu FB, Fuchs C, *et al.* (2003) Major dietary patterns and the risk of colorectal cancer in women. *Arch Intern Med* **163**: 309–314.
- Sieri S, Krogh V, Pala V, *et al.* (2004) Dietary patterns and risk of breast cancer in the ORDET cohort. *Cancer Epidemiol Biomarkers Prev* **13**: 567–572.
- Tseng M, Breslow RA, DeVellis RF, Ziegler RG (2004) Dietary patterns and prostate cancer risk in the National Health and Nutrition Examination Survey Epidemiological Follow-up Study Cohort. *Cancer Epidemiol Biomarkers Prev* **13**: 71–77.

21. Williams DEM, Prevost AT, Whichelow MJ, Cox BD, Day NE, Wareham NJ (2000) A cross-sectional study of dietary patterns with glucose intolerance and other features of the metabolic syndrome. *Br J Nutr* **83**: 257–266.
22. Appel L, Moore TJ, Obarzanek E, for the DASH collaborative research group (1997). A clinical trial of the effects of dietary patterns on blood pressure. *N Engl J Med* **336**: 1117–1124.
23. Tuomilehto J, Lindstrom J, Eriksson JG, *et al.* (2001) Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* **344**: 1343–1350.
24. Martinez ME, Marshall JR, Sechrest L (1998) Invited commentary: factor analysis and the search for objectivity. *Am J Epidemiol* **148**: 17–19.
25. Jacques PF, Tucker KL (2001) Are dietary patterns useful for understanding the role of diet in chronic disease?. *Am J Clin Nutr* **73**: 61–67.
26. Balder HF, Brants HAM, Virtanen M, *et al.* (2003) Common and country-specific dietary patterns in four European cohort studies. *J Nutr* **133**: 4246–4251.
27. Brandt PA van den, Goldbohm RA, van't Veer P, Volovics A, Hermus RJ, Sturmans F (1990) A large-scale prospective cohort study on diet and cancer in the Netherlands. *J Clin Epidemiol* **43**: 285–295.
28. Sieri S, Krogh V, Muti P, *et al.* (2002) Fat and protein intake and subsequent breast cancer risk in postmenopausal women. *Nutr Cancer* **42**: 10–17.
29. Michels KB, Holmberg L, Bergkvist L, Ljung H, Bruce A, Wolk A (2001) Dietary antioxidants vitamins, retinal, and breast cancer incidence in a cohort of Swedish women. *Int J Cancer* **91**: 563–567.
30. Goldbohm RA, Brandt PAVan den, Brants HAM, van't Veer P, Ai M, Sturmans F, Hermus RJJ (1994) Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *Eur J Clin Nutr* **48**: 253–265.
31. Wolk A, Bergström R, Hunter D, *et al.* (1998) A prospective study of association of monounsaturated fat and other types of fat with risk of breast cancer. *Arch Int Med* **158**: 41–45.
32. Mardia KV, Kent JT, Bibby MC (1979) *Multivariate Analysis*, Academic Press., London, UK .
33. Voorrips LE, Brants HA, Kardinaal AF, Hiddink GJ, van den Brandt PA, Goldbohm RA (2002) Intake of conjugated linoleic acid, fat, and other fatty acids in relation to postmenopausal breast cancer: the Netherlands cohort study on diet and cancer. *Am J Clin Nutr* **76**: 873–882.
34. Block G, Patterson B, Subar A (1992) Fruit, vegetables, and cancer prevention: a review of the epidemiological evidence. *Nutr Cancer* **18**: 1–29.
35. Steinmetz KA, Potter JD (1996) Vegetables, fruit, and cancer prevention: a review. *J Am Diet Assoc* **96**: 1027–1039.
36. Dorant E, van den Brandt PA, Goldbohm RA (1995) Allium vegetable consumption, garlic supplement intake, and female breast carcinoma incidence. *Br Cancer Res Treat* **33**: 163–170.
37. Verhoeven DT, Assen N, Goldbohm RA, *et al.* (1997) Vitamins C and E, retinal, beta-carotene and dietary fibre in relation to breast cancer risk: a prospective cohort study. *Br J Cancer* **75**: 149–155.
38. Smith-Warner SA, Spiegelman D, Yaun SS, *et al.* (2001) Intake of fruit and vegetables and risk of breast cancer. A pooled analysis of cohort studies. *J Am Med Assoc* **285**: 769–776.
39. Gaard M, Tretli S, Loken EB (1995) Dietary fat and the risk of breast cancer: a prospective study of 25,892 Norwegian women. *Int J Cancer* **63**: 13–17.
40. Ambrosone CB, Freudenheim JL, Sinha R, *et al.* (1998) Breast cancer risk, meat consumption and N-acetyltransferase (NAT2) genetic polymorphisms. *Int J Cancer* **75**: 825–830.
41. Zheng W, Gustafson DR, Sinha R, *et al.* (1998) Well-done meat intake and the risk of breast cancer. *J Natl Cancer Inst* **90**: 1724–1729.
42. Missmer SA, Smith-Warner SA, Spiegelman D, *et al.* (2002) Meat and dairy food consumption and breast cancer: a pooled analysis of cohort studies. *Int J Epidemiol* **31**: 78–85.
43. La Vecchia C, Negri E, Franceschi S, Decarli A, Giacosa A, Lipworth L (1995) Olive oil, other dietary fats, and the risk of breast cancer (Italy). *Cancer Causes Control* **6**: 545–550.
44. Trichopoulos A, Katsouyanni K, Stuver S, *et al.* (1995) Consumption of olive oil and specific food groups in relation to breast cancer risk in Greece. *J Natl Cancer Inst* **87**: 110–116.
45. Trichopoulos D, Laggiou P (2001) Dietary patterns and mortality. *Br J Nutr* **85**: 133–134.
46. Dixon LB, Balder HF, Virtanen M, *et al.* (2004) Dietary patterns associated with colon and rectal cancers: Results from the DIETSCAN project. *Am J Clin Nutr* **80**: 1003–1011.
47. Osler M, Andreassen H, Heitmann B, *et al.* (2002) Food intake patterns and risk of coronary heart disease: a prospective cohort study examining the use of traditional scoring techniques. *Eur J Clin Nutr* **56**: 568–574.
48. Subar AF, Kipnis V, Troiano RP, *et al.* (2003) Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Clin Epidemiol* **158**: 1–13.
49. Kipnis V, Subar AF, Midthune D, *et al.* (2003) Structure of dietary measurement error: results of the OPEN Biomarker Study. *Am J Epidemiol* **158**: 14–21.
50. 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.
51. Khani BR, Ye W, Terry P, Wolk A (2004) Reproducibility and validity of major dietary patterns among Swedish women assessed with a food-frequency questionnaire. *J Nutr* **134**: 1541–1545.
52. Colditz GA (1995) Fat, estrogens, and the time frame for prevention of breast cancer. *Epidemiology* **6**: 209–211.