



**HAL**  
open science

## **Long-term adherence to healthful and unhealthful plant-based diets and breast cancer risk overall and by hormone receptor and histologic subtypes among postmenopausal females**

Sanam Shah, Yahya Mahamat-Saleh, Wassila Ait-Hadad, Nicholas A. Koemel, Raphaëlle Varraso, Marie Christine Boutron-Ruault, Nasser Laouali

### ► To cite this version:

Sanam Shah, Yahya Mahamat-Saleh, Wassila Ait-Hadad, Nicholas A. Koemel, Raphaëlle Varraso, et al.. Long-term adherence to healthful and unhealthful plant-based diets and breast cancer risk overall and by hormone receptor and histologic subtypes among postmenopausal females. *The American Journal of Clinical Nutrition*, 2023, 117 (3), pp.467-476. <10.1016/j.ajcnut.2022.11.019>. <hal-04069270>

**HAL Id: hal-04069270**

**<https://hal.science/hal-04069270v1>**

Submitted on 31 Mar 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Copyright - All rights reserved

Long-term adherence to healthful and unhealthful plant-based diets and breast cancer risk overall and by hormone receptor and histologic subtypes among postmenopausal women.

Sanam Shah<sup>1</sup>; Yahya Mahamat-Saleh<sup>2</sup>; Wassila Ait-Hadad<sup>3</sup>; Nicholas A. Koemel<sup>4,5</sup>;  
Raphaëlle Varraso<sup>3</sup>; Marie-Christine Boutron-Ruault<sup>1</sup>; Nasser Laouali<sup>1,6,7</sup>

#### Affiliations

<sup>1</sup>Paris-Saclay University, UVSQ, Univ. Paris-Sud, Inserm, Gustave Roussy, “Exposome and Heredity” team, CESP, F-94805, Villejuif, France

<sup>2</sup>International Agency for Research on Cancer (IARC/WHO), 150 cours Albert Thomas, 69372, CEDEX 08, Lyon, France

<sup>3</sup>Paris-Saclay University, UVSQ, Univ. Paris-Sud, Inserm, Gustave Roussy, “Integrative Respiratory Epidemiology” team, CESP, F-94805, Villejuif, France

<sup>4</sup>Charles Perkins Centre, The University of Sydney, Sydney, Australia

<sup>5</sup>Sydney Medical School, The University of Sydney, Sydney, Australia

<sup>6</sup>Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, Massachusetts, USA

<sup>7</sup>Scripps Institution of Oceanography, University of California, San Diego, USA

Email: marie-christine.boutron@imm.fr

\* Marie-Christine Boutron-Ruault and Nasser Laouali contributed equally as the last authors.

#### Funding statement

This research was carried out using data from Inserm's E3N cohort with the support of the MGEN, Institut Gustave Roussy and the “*Ligue contre le Cancer*” for the constitution and maintenance of the E3N cohort. This work has also benefited from State aid managed by the National Research Agency under the program “*Investissement d’avenir*” under the reference ANR-10-COHO-0006 as well as a subsidy from the “*Ministère de l’enseignement supérieur de la recherche et de l’innovation*” for public service charges under the reference n°2103 586016. Sanam Shah is supported by a doctoral funding from *l’Ecole Doctorale de Santé Publique, Ministère de l’enseignement supérieur, de la recherche et de l’innovation*. Nasser Laouali is supported by a research fellowship from the *Fondation de France*.

Data availability: The datasets generated during and/or analyzed for the current study are available from the corresponding author upon reasonable request.

Running head: Plant-based diet and breast cancer

Abbreviations :

BC- Breast cancer

BMI- Body mass index

CI- Confidence interval

DCIS- Ductal carcinoma in situ

E3N- *Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l’Education Nationale*

ER- Estrogen

ER+ - ER-positive

ER- - ER-negative

ER+PR+ - Estrogen-positive Progesterone-positive

ER-PR- - Estrogen-negative Progesterone-negative

FFQ- Food frequency questionnaire

HR- Hazard ratio

hPDI- Healthful plant-based diet index

IDC- Invasive ductal carcinoma

MHT- Menopausal hormone therapy

PR- Progesterone

PR+ - Progesterone-positive

PR- - Progesterone-negative

PDI- Plant-based diet index

SD- Standard deviation

uPDI- Unhealthful plant-based diet index

Registry: The protocol is registered at [clinicaltrials.gov](https://clinicaltrials.gov) as NCT03285230

## Abstract

**Background:** Epidemiological studies assessing the influence of vegetarian diets on breast cancer (BC) risk have produced inconsistent results. Few studies have assessed how the incremental decrease in animal foods and the quality of plant foods are linked with BC.

**Objective:** Disentangle the influence of plant-based diet quality on BC risk among postmenopausal women.

**Design:** 65,574 participants from the *Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale* (E3N) cohort were followed from 1993 to 2014. Incident BC cases were confirmed through pathological reports and classified into subtypes. Cumulative average scores for healthful (hPDI) and unhealthful (uPDI) plant-based diet indices were developed using self-reported dietary intakes at baseline (1993) and follow-up (2005) and divided into quintiles. Cox proportional hazards models were used to estimate adjusted hazard ratios (HR) and 95% confidence intervals (CI).

**Results:** During a mean follow-up of 21 years, 3,968 incident postmenopausal BC cases were identified. There was a non-linear association between adherence to hPDI and BC risk ( $P_{\text{non-linear}} < 0.01$ ). Compared to participants with low adherence to hPDI, those with high adherence had a lower BC risk [HR<sub>Q3 vs. Q1</sub> (95% CI): 0.79 (0.71, 0.87) and HR<sub>Q4 vs. Q1</sub> (95% CI): 0.78 (0.70, 0.86)]. In contrast, higher adherence to uPDI was associated with a linear increase in BC risk [ $P_{\text{non-linear}} = 0.18$ ; HR<sub>Q5 vs. Q1</sub> (95% CI): 1.20 (1.08, 1.33),  $P_{\text{trend}} < 0.01$ ]. Associations were similar according to BC subtypes ( $P_{\text{heterogeneity}} > 0.05$  for all).

**Conclusions:** Long-term adherence to healthful plant foods with some intake of unhealthy plant and animal foods may reduce BC risk with an optimal risk reduction in the moderate

intake range. Adherence to an unhealthful plant-based diet may increase BC risk. These results emphasize the importance of the quality of plant foods for cancer prevention.

**Keywords:** Breast cancer; Plant-based diet quality; Dietary score; Estrogen receptor; Progesterone receptor; Prospective study

## **Introduction**

Breast cancer (BC) is one of the leading global health challenges and, in 2020, accounting for an estimated 2.3 million new cases and 685,000 deaths worldwide (1). It is one of the leading causes of disability-adjusted life years (15.1 million from 1990 to 2015) (2). Cases are predicted to rise in the coming years, particularly postmenopausal BC and estrogen-positive (ER+) tumors (3, 4).

In addition to older age, several factors may contribute to BC risk. Armstrong et al. estimated that the prevalence of genetic mutations accounting for hereditary BC ranged from 0.6% to 36.9% internationally (5), leaving a sizeable unexplained proportion potentially due to modifiable risk factors, including environmental exposures and reproductive and lifestyle factors (6-10). With the increase in life expectancy worldwide, there is a strong interest in identifying modifiable factors involved in the occurrence of BC. As such, research has focused on understanding lifestyle factors, particularly diet.

Plant-based diets have received extensive attention recently due to their benefits for individual health and environmental sustainability (11). Vegetarian diets with a binary characterization in terms of exclusion or not of animal-based foods have long been explored in relation to BC risk, but the results are inconsistent (12). Based on the incremental decrease in animal foods consumption and the notion that the nutritional quality is not consistent across all plant-based foods, Satija et al. have recently proposed healthful and unhealthful plant-based dietary patterns (13). These indices have gained attention for their potential to prevent or manage chronic diseases (13-19). Despite retrospective and prospective studies on the link between plant-based diet indices with BC, findings remain inconsistent (20-25).

Breast cancer is a heterogeneous disease with several subtypes with unique disease progression and response to treatment (26, 27). Estrogen-positive and/or progesterone-positive (PR+) BCs have a substantial advantage from oncologic endocrine therapy in addition to chemotherapy, while the treatment of triple-negative BCs [estrogen-negative (ER-), progesterone-negative (PR-), and human epidermal growth factor receptor 2-negative] depends on chemotherapy alone (28, 29). Even though a potential heterogeneity has been suggested in the associations between intakes of some nutrients (30), food groups (31), and dietary patterns (32, 33) and the risk of specific BC types according to, e.g., ER, PR status, and histology, to our knowledge, only one study investigated a potential heterogeneity in the association between a plant-based diet and BC by ER status (20).

Therefore, we aimed to clarify further the relationship between plant-based diet quality and BC risk overall and by subtypes of BC defined by ER and PR status and histology in participants from the *Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale* (E3N) cohort study. Since associations between BC risk and environmental factors have been found heterogeneous between pre- and postmenopausal BC, the latter being the most frequent, we restricted our analyses to postmenopausal BC in this cohort of middle-aged and elderly participants.

## **Material and Methods**

### **E3N cohort**

The E3N cohort was initiated in 1990 to prospectively investigate the risk factors of common cancers. Participants were 98,995 French women aged 40 to 65 years at inclusion, selected from the health insurance scheme covering workers in the National Education System and

their families (34). The study participants provided written informed consent, and the cohort study received ethical approval from the French National Commission for Computerized Data and Individual Freedom. Participants were enrolled in the cohort through a self-administered questionnaire followed by questionnaires every 2 to 3 years for sociodemographic factors, health conditions, reproductive factors, diet, and other lifestyle factors.

### **Study population**

In the present study, follow-up began on the return date of the first dietary questionnaire for participants who were already menopausal at that time or the date of menopause if it occurred later. Participants contributed person-time until the date of diagnosis of any type of cancer except basal cell carcinoma, the date of the last completed questionnaire, or the date on which the last available follow-up questionnaire was mailed (November 17, 2014), whichever occurred first (35).

Among 74,522 participants who returned the dietary questionnaire sent in 1993, we first excluded those with undefined menopausal status ( $n = 14$ ), those who had never menstruated ( $n = 6$ ), prevalent cancer cases ( $n = 4,709$ ), those with incomplete or absent follow-up information ( $n = 623$ ); we further excluded participants with extreme energy intake values, i.e., the 1<sup>st</sup> and 99<sup>th</sup> percentiles of the energy intake over energy requirement distribution in the population ( $n = 1,364$ ), those with missing BC receptor status data ( $n = 1,309$ ), and those who had not attained menopause at the end of follow-up ( $n = 923$ ). Hence, our final study population included 65,574 postmenopausal participants (**Supplementary Figure 1**).

### **Dietary assessment and plant-based diet indices**

Dietary data were collected at baseline (1993) and follow-up (2005) using a validated self-administered 208-item food frequency questionnaire (FFQ) (36, 37). Participants were asked

about the frequency of consumption for eight eating moments from breakfast to after-dinner snacks over the preceding year. According to the French meal patterns, portion sizes were assessed via photographs and qualitative questions on specific food and drink items. Nutrient and energy intakes were obtained using the Food Composition Database derived from the French Information Center on Food Quality (38).

The following indices were created using a procedure previously described (13): healthful PDI (hPDI), which consisted of healthy plant foods, such as fruit, vegetables, nuts, and legumes, and the unhealthy PDI (uPDI), which consisted of primarily refined/processed foods as listed in **Supplementary Table 1**. First, we created 18 food groups (summing up the grams of consumed food items). Next, participants were categorized according to the quintile distribution (or a non-consumer category plus quartiles among consumers) of their food intake for each food group. For each participant, values of 1 to 5 or reverse were assigned to each category based on the positive or negative association with the index; animal foods were reverse scored. The sum of the values for each food parameter led to the final hPDI and uPDI scores, which ranged from 18 to 90 (lowest to highest adherence).

These scores have been developed in the E3N cohort using only baseline dietary data (19); however, in the present study, the scores from baseline (1993) and follow-up (2005) were averaged to better capture the long-term adherence and changes in diet over the long follow-up. Baseline scores were used if participants were censored before the follow-up dietary questionnaire (2005). The baseline and cumulative average scores were used for the final analysis for 12,689 and 52,885 participants, respectively. The observed median values were 56 (range 27-79) for hPDI and 52.5 (range 27-78) for uPDI.

### **Incident breast cancer ascertainment**

All potential cases of BC were identified through baseline and follow-up questionnaires (3<sup>rd</sup> to 11<sup>th</sup>) which inquired about cancer occurrence, contact details of participants' physicians, and permission to contact them. A few BC cases were further identified from insurance files and death certificates. Tumor characteristics of hormonal receptor status ER and PR and histology (ductal, lobular, other) were extracted using original clinical and pathology reports. The pathology reports were obtained for 95% of incident BC cases.

### **Covariates**

The covariates were selected a priori based on literature evidence for their potential association with the exposure and/or the outcome. Education, physical activity (MET-h/week), and smoking status were self-reported at baseline. Breastfeeding and lifetime use of oral contraception and menopausal hormone therapy (MHT) were assessed from the baseline and follow-up questionnaires. Age at menarche, history of benign breast disease, parity, and family history of BC were self-reported at baseline. Mammography performed in the previous follow-up cycle was captured from each follow-up questionnaire. Body mass index (BMI) was assigned according to the value reported at baseline; self-reported height and weight were used to calculate BMI, defined as the weight (kg) divided by squared height (m<sup>2</sup>). In the cohort, self-reported anthropometry is considered reliable from a validation study (39). Lastly, alcohol consumption (grams/day) and energy intake were calculated from the E3N FFQ.

### **Statistical analysis**

Baseline characteristics overall and according to hPDI and uPDI quintiles were described using means and standard deviations (SD) for continuous variables and frequencies for categorical variables. Hazard ratios (HR) and 95% confidence intervals (CI) of the BC risk were estimated using multivariable-adjusted Cox proportional hazards regression with age as

the time scale (entry time defined as the age at the baseline questionnaire or age at menopause, whichever was maximum). The proportional hazards assumption was tested graphically using Schoenfeld Residuals, and no major violations were observed (40).

All diet indices were analyzed in three ways. We first tested for non-departure from a linear association and provided a graphical representation using restricted cubic splines (41). The splines analyses were fitted with the fully adjusted model, and five knots were placed at the 5<sup>th</sup>, 27.5<sup>th</sup>, 50<sup>th</sup>, 72.5<sup>th</sup>, and 95<sup>th</sup> percentiles. Second, when a linear association was confirmed, we assessed the HR of BC risk for a 1-SD increase in the diet index. Lastly, all diet indices were categorized into quintiles, and the first quintile group was considered the reference category to assess whether higher categories were associated with BC risk.

All models were stratified by 5-year birth cohorts, and three sets of multivariable Cox models were built by adding covariates: Model 1 was age-adjusted. Model 2 additionally included physical activity (continuous), educational level (undergraduate or less, graduate, postgraduate or more), smoking status (current, former, non-smoker), family history of BC (yes, no), age at menarche (continuous), age at first childbirth (nulliparous, <30 years,  $\geq$  30 years), ever breastfeeding (yes, no), ever use of MHT (yes, no), ever use of the contraceptive pill (yes, no), past history of benign breast disease (yes, no), and mammography in the last follow-up cycle (yes, no). The final Model 3 also included potential diet and cancer association mediators, BMI (continuous), energy intake (excluding alcohol, continuous), and alcohol (continuous). The P-value for linear trend was estimated in the models using the median score in each quintile. Subtypes of BC, characterized by hormone receptors and histology, were studied in separate Cox models. We used the Q statistic to test the homogeneity of the results between the receptor subtypes and histology (42).

In addition to the above-described primary analyses, we tested for effect modification of the overall associations by BMI (< 20, 20-24.99, > 25 kg/m<sup>2</sup>), use of MHT (yes, no), and physical activity (above or below median physical activity level, i.e., < 37.97, ≥ 37.97 MET-h/week) as suggested by previous studies (20, 43). Due to a small number of cases, the stratified analyses were not performed for the subtypes.

Missing observations were < 5% for all variables except for ever breastfed and therefore were imputed to the median (for continuous variables) or mode (for categorical variables). For ever breastfed, a ‘missing’ category was created to maintain the same number of participants in the analyses. All tests of statistical significance were two-sided, with statistical significance set at  $P < 0.05$ . Statistical analyses were performed using the SAS software (version 9.4 SAS Institute, Cary, NC).

### **Sensitivity analyses**

Several secondary analyses were conducted to assess the robustness of the findings. First, missing values in the covariates were imputed with multiple imputation to overcome any systematic differences between the participants with complete and missing information. Second, the participants with BC diagnosed in the first five years of follow-up were excluded to overcome the influence of any reverse causation bias due to changes in diet after diagnosis of BC. Third, we analyzed the association of the overall PDI with the risk of BC. Fourth, we used the baseline and the time-dependent approaches to analyze dietary measurements across the follow-up. The cumulative average approach differs from the time-dependent approach in that the incident BC diagnosed between 2005 and 2014 was related to the average dietary intake reported on the 1993 and 2005 questionnaires instead of the dietary intake on the 2005 questionnaire only. Fifth, as we had too few cases of ductal carcinoma in situ (DCIS) (n=158) to conduct separate analyses, we performed additional analyses based only on invasive ductal

carcinoma (IDC) cases. Lastly, we excluded participants with energy limits <500 or >3500 kcal/day as proposed by Willett instead of the 1<sup>st</sup> and 99<sup>th</sup> percentiles energy distribution.

## **Results**

### **Baseline characteristics**

Three thousand nine hundred and sixty-eight incident BC cases were diagnosed among 65,574 postmenopausal participants over approximately 21 years of follow-up. The average age of participants at baseline was 52.9 years (SD 6.7).

Participants in the highest quintile of cumulative average hPDI had lower BMI. They were more likely to have lower levels of energy and alcohol intakes, to report lower use of the contraceptive pill, and to be former smokers compared to those in the lowest quintile. They were also more likely to report higher levels of physical activity, a history of mammography, a history of benign breast disease, and ever use of MHT compared to participants in the lowest quintile (**Table 1**). The consumption in grams/day of the food groups according to quintiles of cumulative average hPDI is presented in **Supplementary Table 2**. Participants in the highest quintile of cumulative average uPDI reported lower levels of energy and alcohol intakes and lower physical activity, were more likely to have lower BMI and lower proportion of excess weight, and were less likely to be current smokers, to report ever use of the contraceptive pill, to have a history of benign breast disease, and to have recent mammography compared to participants in the lowest quintile (**Supplementary Table 3**).

### **Healthful and unhealthful plant-based diet indices and BC risk overall**

**Table 2 and Figures 1-3** present the associations of cumulative average hPDI and uPDI with postmenopausal BC risk. After considering known BC risk factors, spline analysis showed a significant departure from linearity for the association between hPDI and BC risk ( $P_{\text{non-linear}} < 0.01$ ) (**Figure 1**). Participants with hPDI score intervals of approximately 52 to 62 had the lowest BC risk. Similarly, when the hPDI was modeled as a categorical variable, those in hPDI quintiles 3 and 4 had the lowest BC risk, 21% and 22%, respectively [fully-adjusted model (Model 3),  $HR_{Q3}$  (95% CI): 0.79 (0.71, 0.87) and  $HR_{Q4}$  (95% CI): 0.78 (0.70, 0.86)] compared to those in the first quintile. Participants in the highest quintile 5 had a 14% lower risk of BC [ $HR_{Q5 \text{ vs. } Q1}$  (95% CI): 0.86 (0.77, 0.95)].

In contrast, for cumulative average uPDI, the spline analysis was consistent with a linear relation ( $P_{\text{non-linear}} = 0.18$ ) (**Figure 2**). Modeled as a continuous variable, uPDI resulted in a 4% higher risk of BC [Model 3,  $HR_{1\text{-SD increase}}$  (95% CI): 1.04 (1.01, 1.08)]. When quintile groups of uPDI were considered, a 20% higher risk for the highest vs. lowest quintile was observed [Model 3,  $HR_{Q5 \text{ vs. } Q1}$  (95% CI): 1.20 (1.08, 1.33),  $P_{\text{trend}} < 0.01$ ]. Since the HRs changed substantially from Model 2 to 3, we adjusted on energy intake, alcohol, and BMI individually to assess which variables could be driving the associations. We observed the following changes: adjusting for BMI,  $HR_{Q5 \text{ vs. } Q1}$  (95% CI): 1.14 (1.03, 1.26), for energy intake,  $HR_{Q5 \text{ vs. } Q1}$  (95% CI): 1.17 (1.05, 0.30), and for alcohol,  $HR_{Q5 \text{ vs. } Q1}$  (95% CI): 1.13 (1.03, 1.25). The greatest variation in HR occurred when adjusting for energy intake, suggesting that this variable could be an important confounder or mediator of the association between uPDI and BC risk.

### **Healthful and unhealthful plant-based diet indices according to hormone receptor and histologic subtypes**

When investigating the association of cumulative average hPDI according to BC subtypes, we did not observe heterogeneity by either receptor status or histology ( $P_{\text{heterogeneity}} > 0.05$  for all) (**Supplementary Tables 4-7**). For example, in the fully-adjusted model, when we compared the risks associated with quintile 5 of hPDI for PR+ and PR-, we observed 13% and 16% lower risks, respectively [HR<sub>Q5 vs. Q1</sub> (95% CI): 0.87 (0.76, 0.99) and 0.84 (0.70, 1.00);  $P_{\text{heterogeneity}} = 0.76$ ] (**Figure 3**). For ER+ BC, participants in the highest quintile of hPDI had a 15% lower BC risk [HR<sub>Q5 vs. Q1</sub> (95% CI): 0.85 (0.76, 0.95)] compared with participants in the lowest quintile. For ER- BC, there was a lower risk for the first two quintiles and a null association for quintile 5 [HR<sub>Q5 vs. Q1</sub> (95% CI): 0.88 (0.69, 1.12)] (**Supplementary Table 4**). The HR<sub>SQ5 vs. Q1</sub> (95% CI) for ER+PR+ and ER-PR- BC were 0.86 (0.75, 0.99) and 0.87 (0.67, 1.14), respectively (**Supplementary Table 6**). Lastly, for ductal and lobular BC, the HR<sub>SQ5 vs. Q1</sub> (95% CI) were 0.84 (0.74, 0.95) and 0.79 (0.62, 1.02), respectively (**Supplementary Table 7**).

In subgroup analyses for cumulative average uPDI, in the fully-adjusted model, participants in the highest intake quintile had a 20% higher risk of ER+ BC, HR<sub>Q5 vs. Q1</sub> (95% CI): 1.20 (1.07, 1.35,  $P_{\text{trend}} < 0.01$ ), albeit we did not observe heterogeneity by either receptor status or histology ( $P_{\text{heterogeneity}} > 0.05$  for all) (**Figure 3**). Corresponding HR<sub>Q5 vs. Q1</sub> (95% CI) for ER- BC was 1.20 (0.94, 1.53),  $P_{\text{trend}} = 0.45$ ) (**Supplementary Table 4**). There was an 18% and 24% higher risk of PR+ and PR- BC, HR<sub>Q5 vs. Q1</sub> (95% CI): 1.18 (1.03, 1.34),  $P_{\text{trend}} = 0.02$  and 1.24 (1.05, 1.47),  $P_{\text{trend}} = 0.02$ , respectively (**Supplementary Table 5**). Corresponding HR<sub>SQ5 vs. Q1</sub> (95% CI) for ER+PR+ and ER-PR- BC were 1.17 (1.02, 1.34),  $P_{\text{trend}} = 0.02$  and 1.19 (0.91, 1.55),  $P_{\text{trend}} = 0.57$ , respectively (**Supplementary Table 6**). Lastly, there were positive associations for both ductal and lobular BCs, HR<sub>Q5 vs. Q1</sub> (95% CI): 1.17 (1.03, 1.32),  $P_{\text{trend}} = 0.02$ , and 1.55 (1.20, 1.99),  $P_{\text{trend}} < 0.01$ , respectively (**Supplementary Table 7**).

## Stratified analyses

There were no interactions between cumulative average hPDI and uPDI and BMI, MHT use, or physical activity with respect to BC risk ( $P_{\text{interaction}} > 0.05$  for all) (**Supplementary Tables 8-10 and Supplementary Figure 2**).

## Sensitivity analyses

In sensitivity analyses using multiple imputations to handle missing covariates, a similar pattern of results was obtained compared to those of the primary analysis (results not tabulated). In addition, results were unchanged when excluding participants diagnosed in the first five years of follow-up, suggesting that reverse causation was unlikely to explain the findings (results not tabulated).

We found evidence for a non-linear association between the cumulative average overall PDI and BC risk ( $P_{\text{non-linear}} = 0.03$ ). Considering quintiles of PDI in the fully-adjusted model, there was a lower risk of BC from quintiles 2 to 4 [HR<sub>Q2</sub> (95% CI): 0.83 (0.75, 0.92) to HR<sub>Q4</sub> (95% CI): 0.84 (0.76, 0.93)] which attenuated in quintile 5 [HR (95% CI): 0.94 (0.85, 1.04)].

Furthermore, when using scores derived by FFQ at baseline, in Model 3, there was a linear association between PDI ( $P_{\text{non-linear}} = 0.51$ ), hPDI ( $P_{\text{non-linear}} = 0.86$ ), and uPDI ( $P_{\text{non-linear}} = 0.11$ ) and BC risk. Considering a 1-SD increase in hPDI in Model 3, there was a 3% lower risk of BC [HR (95% CI): 0.97 (0.93, 1.00)]. Considering quintiles of baseline hPDI, there was an 11% lower risk of BC [HR (95% CI): 0.89 (0.81, 0.99),  $P_{\text{trend}} < 0.01$ ] in Model 1, while full adjustment attenuated the association, HR (95% CI): 0.92 (0.83, 1.02),  $P_{\text{trend}} = 0.06$ . Conversely, for baseline uPDI, there was a 4% increase in risk for a 1-SD increase [HR (95% CI): 1.04 (1.01, 1.08)]. As quintiles of uPDI, in Model 3, there was a 13% higher risk [HR (95% CI): 1.13 (1.02, 1.26),  $P_{\text{trend}} = 0.02$ ]. We found no association between baseline overall PDI and BC risk (results not tabulated).

When considering the time-dependent indexes in the fully adjusted model, for each 1-SD increase, the HRs (95% CI) for overall BC risk were as follows: PDI, 0.977 (0.969, 0.985); hPDI, 0.991 (0.983, 0.999); and uPDI, 0.997 (0.988, 1.005). For the quintiles group analysis, the HR (95% CI) for overall BC risk were as follows: PDI, HR<sub>Q5 vs. Q1</sub> (95% CI): 0.94 (0.91, 0.96); hPDI, HR<sub>Q5 vs. Q1</sub> (95% CI): 0.97 (0.94, 1.00); and uPDI, HR<sub>Q5 vs. Q1</sub> (95% CI): 0.99 (0.96, 1.01) (results not tabulated).

The HRs CI for IDC was not materially different from our primary analysis; for cumulative average hPDI and uPDI, in the fully-adjusted model, the HR<sub>Q5 vs. Q1</sub> (95% CI) were 0.84 (0.74, 0.95) and 0.18 (1.04, 1.34),  $P_{\text{trend}} < 0.01$ , respectively (results not tabulated).

Lastly, when we used energy limits proposed by Willett, the results did not change materially, and the direction of the association remained. For cumulative average hPDI and uPDI in the fully-adjusted model, the HR<sub>Q5 vs. Q1</sub> (95% CI) were 0.85 (0.77, 0.94) and 1.16 (1.05, 1.28),  $P_{\text{trend}} = 0.02$ , respectively.

## **Discussion**

In the French E3N prospective cohort study, a non-linear association was observed between adherence to hPDI and BC risk, with the lowest risk for moderate adherence. In contrast, higher adherence to uPDI was positively associated with BC risk. In addition, we found no heterogeneity between adherence to these plant-based diet indices and BC risk across tumor subtypes. These results suggest that hPDI could have a preventive influence and uPDI a promotive influence on all types of BC, thus acting independently of hormonal mechanisms and reflecting the importance of plant-based diet quality in BC prevention in postmenopausal women.

Our findings of a differential pattern of association between a plant-based diet and postmenopausal BC risk by the quality of plant foods were observed in previous studies. In a cohort study of more than 150,000 participants from the Nurses' Health Studies (NHS), an inverse association was observed with hPDI. However, they did not evidence any increased risk of overall BC with uPDI (20). Moreover, in a case-control study on 350 cases and 700 controls from Iran, there were lower odds of BC in postmenopausal participants with the greatest adherence to hPDI, whereas the greatest adherence to uPDI was associated with higher odds (21), although in another case-control study on 412 cases and 456 controls, hPDI was associated with lower odds, but no association was observed with uPDI (23). Furthermore, in a Mediterranean cohort study of 10,812 participants, neither the healthful nor the unhealthful provegetarian patterns were associated with postmenopausal BC risk (25).

Although a non-linear association between high-quality plant foods such as fruit and vegetables and BC risk has not been reported so far, a non-linear association between dietary folate intake and BC risk in the EPIC cohort was suggested (44). However, such an association has already been reported for colorectal cancer. For example, a meta-analysis of 19 prospective studies conducted by Aune et al. reported a non-linear association between fruit and vegetable intake and colorectal cancer risk, with the greatest risk reduction in the lower intake range (45). Furthermore, a meta-analysis by Schwingshackl reported non-linear associations between vegetables, fruit, and nuts with all-cause mortality (46).

We found no heterogeneity across BC subtypes in our study. The risk of BC by receptor and histological subtypes has not been extensively studied. As regards the association between dietary patterns and ductal and lobular carcinoma, studies support the findings of an inverse risk between the Mediterranean diet and risk of IDC and invasive lobular carcinoma (ILC) BC and a positive association between the Western diet and IDC and ILC BC (47); however,

literature is scarce for associations between BC histological types and plant-based diets. The NHS study reported heterogeneity by ER receptor status, with the strongest association between hPDI and ER- tumors (20).

The discrepancies in findings of these studies could partly be attributed to varying sample sizes, differences in the development of scores using 18- (20, 25) or 16- food group constructs (21), dietary measurement errors, the duration of follow-up (which varies between 115,802 and over 4,841,083 person-years of follow-up for prospective studies), as well as the non-reporting in some studies of the potential deviation from the linearity of association.

Several potential biological mechanisms support a protective role of hPDI towards BC risk.

The healthful plant-predominant dietary pattern is characterized by high fiber, fruit, vegetables, nuts, legumes, and low red and processed meat. Hence, abundant antioxidants, vitamins, and fiber could have anti-inflammatory and antiproliferative activities (48-50). Furthermore, there is experimental evidence that polyphenols in fruit and vegetables inhibit the proliferation of DCIS BC (51). High fiber also reduces the total body pool of estrogen, thus helping reduce BC risk (52). Similarly, phytoestrogens which are plant-derived estrogen-like compounds, have anticarcinogenic effects of their own and via lowering the amount of circulating estrogen; meta-analyses have reported that dietary lignans, a class of phytoestrogen found in vegetables, fruit, whole grains, green tea, and oilseeds, were associated with lower BC risk in postmenopausal participants (53, 54). In France, a high dietary intake of plant lignans and enterolignans was associated with lower ER+ and PR+ postmenopausal BC risk (55). On the other hand, less healthful plant-based foods are poor in fiber and micronutrients and laden with carbohydrates with tumor-promoting potential via effects on circulating insulin and insulin-like growth factor-1 (56, 57). Trans-fatty acids in industrially processed salty foods, sweets, and other packaged foods influence systemic inflammation, visceral adiposity, body weight, and insulin resistance, increasing BC risk (58-

60). Furthermore, polyunsaturated fatty acid metabolites reduce the binding of estrogen to serum-binding proteins increasing circulating levels and activating breast cell growth (61). More evidence on the effects of bioactive metabolites of various compounds could shed light on these findings.

This study has some important strengths. First, the prospective nature of the study design precluded recall and selection biases with high retention over a long follow-up period. Second, the large sample size provided statistical power to detect associations. Third, the large number of incident cases allowed us to analyze by receptor and histological subtypes. Fourth, validated FFQs were used to collect dietary intake data before BC diagnosis. We used the cumulative average diet scores, thus accounting for possible changes in dietary habits during the follow-up period and better reflecting long-term dietary adherence. Fifth, excluding participants with BC diagnosed in the first five years of follow-up, did not change our results, suggesting that reverse causation was unlikely to explain our findings. Lastly, when we applied the Willett energy exclusion criteria, although it reduced the study population and the study power, the results were somewhat stronger and in favor of a true association.

However, this study was met with some limitations. First, although we adjusted for most known confounding factors of BC, there could still be some bias from unmeasured confounders due to the study's observational nature. Second, self-reported dietary intake was used as exposure information. Although the FFQ has been validated, some degree of non-differential exposure misclassification is likely, which is usually considered to bias the results towards the null in a prospective setting such as ours, although a recent paper suggests that there are possibilities of opposite bias (62). Lastly, our highly educated participants might be more health-conscious. They may not be a representative sample of the general French

population, limiting our results' external validity. However, considering the dietary variations in the general population, there is a higher possibility of finding stronger associations between these indices and the risk of BC.

## **Conclusions**

In conclusion, the results showed a non-linear risk reduction between hPDI and BC, with the lowest BC risk for moderate adherence and a linear risk increase between uPDI and BC among postmenopausal women. These findings have implications for all subtypes of BC and support the notion of the importance of the quality of plant-based foods when consuming plant-predominant diets with a balance of animal foods. Additional research is needed to confirm these findings and better understand the underlying mechanisms involved in these associations. Lastly, more studies are required to determine the associations in relation to premenopausal BC, as there are different mechanisms for hormone-related breast cancer in pre- and postmenopausal women. Because premenopausal BCs are less frequent than postmenopausal BCs, a collaborative study of several cohorts would be an ideal setting.

## Additional information

### Acknowledgements

The research was carried out using data from INSERM (French National Institutes for Health and Medical Research) E3N cohort, which was established and maintained with the support of the *Mutuelle Générale de l'Éducation Nationale* (MGEN), Gustave Roussy, and the French League against Cancer (LNCC). E3N-E4N is also supported by the French National Research Agency (ANR) under the Investment for the Future Program (PIA) (ANR-10-COHO-0006) and by the French Ministry of Higher Education, Research and Innovation (subsidy for public service charges n°2103 586016). The authors are indebted to all participants for their continued participation. They are also grateful to all members of the E3N study group.

### Conflict of interest

The authors declare no duality of interest associated with this manuscript.

### Author contributions

SS, MCBR, and NL conceived and designed the study. MCBR and NL contributed equally as the last authors. SS performed the statistical analysis and drafted the original manuscript. All authors contributed to the interpretation of data discussed in the manuscript, revised it, and approved its final version to be published. NL is the guarantor of this work and, as such, has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

### Disclaimer

Where authors are identified as personnel of the International Agency for Research on Cancer/ World Health Organization, the authors alone are responsible for the views expressed

in this article, and they do not necessarily represent the decisions, policies, or views of the International Agency for Research on Cancer/ World Health Organization.

## References

1. Lei S, Zheng R, Zhang S, Wang S, Chen R, Sun K, Zeng H, Zhou J, Wei W. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. *Cancer Commun (Lond)* 2021;41(11):1183-94. doi: 10.1002/cac2.12207.
2. Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, Dicker DJ, Chimed-Orchir O, Dandona R, Dandona L, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 2017;3(4):524-48. doi: 10.1001/jamaoncol.2016.5688.
3. Heer E, Harper A, Escandor N, Sung H, McCormack V, Fidler-Benaoudia MM. Global burden and trends in premenopausal and postmenopausal breast cancer: a population-based study. *Lancet Glob Health* 2020;8(8):e1027-e37. doi: 10.1016/S2214-109X(20)30215-1.
4. Rosenberg PS, Barker KA, Anderson WF. Estrogen Receptor Status and the Future Burden of Invasive and In Situ Breast Cancers in the United States. *J Natl Cancer Inst* 2015;107(9). doi: 10.1093/jnci/djv159.
5. Armstrong N, Ryder S, Forbes C, Ross J, Quek RG. A systematic review of the international prevalence of BRCA mutation in breast cancer. *Clin Epidemiol* 2019;11:543-61. doi: 10.2147/CLEP.S206949.
6. Momenimovahed Z, Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer (Dove Med Press)* 2019;11:151-64. doi: 10.2147/BCTT.S176070.
7. Bernier MO, Plu-Bureau G, Bossard N, Ayzac L, Thalabard JC. Breastfeeding and risk of breast cancer: a metaanalysis of published studies. *Hum Reprod Update* 2000;6(4):374-86. doi: 10.1093/humupd/6.4.374.
8. McCormack VA, Boffetta P. Today's lifestyles, tomorrow's cancers: trends in lifestyle risk factors for cancer in low- and middle-income countries. *Ann Oncol* 2011;22(11):2349-57. doi: 10.1093/annonc/mdq763.
9. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;15(6):1159-69. doi: 10.1158/1055-9965.EPI-06-0034.
10. Kotepui M. Diet and risk of breast cancer. *Contemp Oncol (Pozn)* 2016;20(1):13-9. doi: 10.5114/wo.2014.40560.
11. Fresan U, Sabate J. Vegetarian Diets: Planetary Health and Its Alignment with Human Health. *Adv Nutr* 2019;10(Suppl\_4):S380-S8. doi: 10.1093/advances/nmz019.
12. Godos J, Bella F, Sciacca S, Galvano F, Grosso G. Vegetarianism and breast, colorectal and prostate cancer risk: an overview and meta-analysis of cohort studies. *J Hum Nutr Diet* 2017;30(3):349-59. doi: 10.1111/jhn.12426.
13. Satija A, Bhupathiraju SN, Rimm EB, Spiegelman D, Chiuve SE, Borgi L, Willett WC, Manson JE, Sun Q, Hu FB. Plant-Based Dietary Patterns and Incidence of Type 2 Diabetes in US Men and Women: Results from Three Prospective Cohort Studies. *PLoS Med* 2016;13(6):e1002039. doi: 10.1371/journal.pmed.1002039.
14. Satija A, Bhupathiraju SN, Spiegelman D, Chiuve SE, Manson JE, Willett W, Rexrode KM, Rimm EB, Hu FB. Healthful and Unhealthful Plant-Based Diets

- and the Risk of Coronary Heart Disease in U.S. Adults. *J Am Coll Cardiol* 2017;70(4):411-22. doi: 10.1016/j.jacc.2017.05.047.
15. Kouvari M, Tsiampalis T, Chrysohoou C, Georgousopoulou E, Skoumas J, Mantzoros CS, Pitsavos CS, Panagiotakos DB. Quality of plant-based diets in relation to 10-year cardiovascular disease risk: the ATTICA cohort study. *Eur J Nutr* 2022. doi: 10.1007/s00394-022-02831-0.
  16. Song S, Lee K, Park S, Shin N, Kim H, Kim J. Association between Unhealthy Plant-Based Diets and Possible Risk of Dyslipidemia. *Nutrients* 2021;13(12). doi: 10.3390/nu13124334.
  17. Chen Z, Qian F, Liu G, Li M, Voortman T, Tobias DK, Ley SH, Bhupathiraju SN, Li LJ, Chavarro JE, et al. Prepregnancy plant-based diets and the risk of gestational diabetes mellitus: a prospective cohort study of 14,926 women. *Am J Clin Nutr* 2021. doi: 10.1093/ajcn/nqab275.
  18. Kim J, Kim H, Giovannucci EL. Quality of plant-based diets and risk of hypertension: a Korean genome and examination study. *Eur J Nutr* 2021;60(7):3841-51. doi: 10.1007/s00394-021-02559-3.
  19. Laouali N, Shah S, MacDonald CJ, Mahamat-Saleh Y, El Fatouhi D, Mancini F, Fagherazzi G, Boutron-Ruault MC. BMI in the Associations of Plant-Based Diets with Type 2 Diabetes and Hypertension Risks in Women: The E3N Prospective Cohort Study. *J Nutr* 2021;151(9):2731-40. doi: 10.1093/jn/nxab158.
  20. Romanos-Nanclares A, Willett WC, Rosner BA, Collins LC, Hu FB, Toledo E, Eliassen AH. Healthful and Unhealthy Plant-Based Diets and Risk of Breast Cancer in U.S. Women: Results from the Nurses' Health Studies. *Cancer Epidemiol Biomarkers Prev* 2021;30(10):1921-31. doi: 10.1158/1055-9965.EPI-21-0352.
  21. Rigi S, Mousavi SM, Benisi-Kohansal S, Azadbakht L, Esmailzadeh A. The association between plant-based dietary patterns and risk of breast cancer: a case-control study. *Sci Rep* 2021;11(1):3391. doi: 10.1038/s41598-021-82659-6.
  22. Kane-Diallo A, Srour B, Sellem L, Deschasaux M, Latino-Martel P, Hercberg S, Galan P, Fassier P, Gueraud F, Pierre FH, et al. Association between a pro plant-based dietary score and cancer risk in the prospective NutriNet-sante cohort. *Int J Cancer* 2018;143(9):2168-76. doi: 10.1002/ijc.31593.
  23. Sasanfar B, Toorang F, Booyani Z, Vassalami F, Mohebbi E, Azadbakht L, Zendehtdel K. Adherence to plant-based dietary pattern and risk of breast cancer among Iranian women. *Eur J Clin Nutr* 2021;75(11):1578-87. doi: 10.1038/s41430-021-00869-7.
  24. Payandeh N, Shahinfar H, Amini MR, Jafari A, Safabakhsh M, Imani H, Shab-Bidar S. The Lack of Association between Plant-Based Dietary Pattern and Breast Cancer: a Hospital-Based Case-Control Study. *Clin Nutr Res* 2021;10(2):115-26. doi: 10.7762/cnr.2021.10.2.115.
  25. Romanos-Nanclares A, Toledo E, Sanchez-Bayona R, Sanchez-Quesada C, Martinez-Gonzalez MA, Gea A. Healthful and unhealthy plant-based food patterns and the incidence of breast cancer: Results from a Mediterranean cohort. *Nutrition* 2020;79-80:110884. doi: 10.1016/j.nut.2020.110884.
  26. Polyak K. Heterogeneity in breast cancer. *J Clin Invest* 2011;121(10):3786-8. doi: 10.1172/JCI60534.
  27. Tamimi RM, Baer HJ, Marotti J, Galan M, Galaburda L, Fu Y, Deitz AC, Connolly JL, Schnitt SJ, Colditz GA, et al. Comparison of molecular

- phenotypes of ductal carcinoma in situ and invasive breast cancer. *Breast Cancer Res* 2008;10(4):R67. doi: 10.1186/bcr2128.
28. Ignatiadis M, Sotiriou C. Luminal breast cancer: from biology to treatment. *Nat Rev Clin Oncol* 2013;10(9):494-506. doi: 10.1038/nrclinonc.2013.124.
  29. Griffiths CL, Olin JL. Triple negative breast cancer: a brief review of its characteristics and treatment options. *J Pharm Pract* 2012;25(3):319-23. doi: 10.1177/0897190012442062.
  30. Fung TT, Hu FB, Hankinson SE, Willett WC, Holmes MD. Low-carbohydrate diets, dietary approaches to stop hypertension-style diets, and the risk of postmenopausal breast cancer. *Am J Epidemiol* 2011;174(6):652-60. doi: 10.1093/aje/kwr148.
  31. Wu Y, Huang R, Wang M, Bernstein L, Bethea TN, Chen C, Chen Y, Eliassen AH, Freedman ND, Gaudet MM, et al. Dairy foods, calcium, and risk of breast cancer overall and for subtypes defined by estrogen receptor status: a pooled analysis of 21 cohort studies. *Am J Clin Nutr* 2021;114(2):450-61. doi: 10.1093/ajcn/nqab097.
  32. Kang JH, Peng C, Rhee JJ, Farvid MS, Willett WC, Hu FB, Rosner BA, Tamimi R, Eliassen AH. Prospective study of a diabetes risk reduction diet and the risk of breast cancer. *Am J Clin Nutr* 2020;112(6):1492-503. doi: 10.1093/ajcn/nqaa268.
  33. Hayati Z, Montazeri V, Shivappa N, Hebert JR, Pirouzpanah S. The association between the inflammatory potential of diet and the risk of histopathological and molecular subtypes of breast cancer in northwestern Iran: Results from the Breast Cancer Risk and Lifestyle study. *Cancer* 2022;128(12):2298-312. doi: 10.1002/cncr.34183.
  34. Clavel-Chapelon F. Cohort Profile: The French E3N Cohort Study. *International Journal of Epidemiology* 2015;44(3):801-9. doi: 10.1093/ije/dyu184.
  35. Cottet V, Touvier M, Fournier A, Touillaud MS, Lafay L, Clavel-Chapelon F, Boutron-Ruault M-C. Postmenopausal breast cancer risk and dietary patterns in the E3N-EPIC prospective cohort study. *Am J Epidemiol* 2009;170(10):1257-67. doi: 10.1093/aje/kwp257.
  36. Lucas F, Niravong M, Villemainot S, Kaaks R, Clavel-Chapelon F. Estimation of food portion size using photographs: validity, strengths, weaknesses and recommendations. *Journal of Human Nutrition and Dietetics* 1995;8(1):65-74. doi: 10.1111/j.1365-277x.1995.tb00296.x.
  37. van Liere M. Relative validity and reproducibility of a French dietary history questionnaire. *International Journal of Epidemiology* 1997;26(90001):128S-36. doi: 10.1093/ije/26.suppl\_1.S128.
  38. Jean-Claude. F, J. I-R, C. T, M. F. Répertoire général des aliments : table de composition = Composition tables. Paris (FRA) 1995.
  39. Tehard B, van Liere MJ, Com Nougue C, Clavel-Chapelon F. Anthropometric measurements and body silhouette of women: validity and perception. *J Am Diet Assoc* 2002;102(12):1779-84. doi: 10.1016/s0002-8223(02)90381-0.
  40. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika* 1982;69(1):239-41. doi: 10.1093/biomet/69.1.239.
  41. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*: Springer, 2013.
  42. Allison PD. *Survival Analysis Using SAS®: A Practical Guide, Second Edition*. Cary, NC: SAS Institute Inc. , 2010.

43. Velie EM, Schairer C, Flood A, He J-P, Khattree R, Schatzkin A. Empirically derived dietary patterns and risk of postmenopausal breast cancer in a large prospective cohort study. *The American Journal of Clinical Nutrition* 2005;82(6):1308-19. doi: 10.1093/ajcn/82.6.1308.
44. Van Puyvelde H, Papadimitriou N, Clasen J, Muller D, Biessy C, Ferrari P, Halkjaer J, Overvad K, Tjonneland A, Fortner RT, et al. Dietary Methyl-Group Donor Intake and Breast Cancer Risk in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Nutrients* 2021;13(6). doi: 10.3390/nu13061843.
45. Aune D, Lau R, Chan DS, Vieira R, Greenwood DC, Kampman E, Norat T. Nonlinear reduction in risk for colorectal cancer by fruit and vegetable intake based on meta-analysis of prospective studies. *Gastroenterology* 2011;141(1):106-18. doi: 10.1053/j.gastro.2011.04.013.
46. Schwingshackl L, Schwedhelm C, Hoffmann G, Lampousi AM, Knuppel S, Iqbal K, Bechthold A, Schlesinger S, Boeing H. Food groups and risk of all-cause mortality: a systematic review and meta-analysis of prospective studies. *Am J Clin Nutr* 2017;105(6):1462-73. doi: 10.3945/ajcn.117.153148.
47. Dianatinasab M, Rezaian M, HaghghatNezad E, Bagheri-Hosseiniabadi Z, Amanat S, Rezaeian S, Masoudi A, Ghiasvand R. Dietary Patterns and Risk of Invasive Ductal and Lobular Breast Carcinomas: A Systematic Review and Meta-analysis. *Clin Breast Cancer* 2020;20(4):e516-e28. doi: 10.1016/j.clbc.2020.03.007.
48. Samtiya M, Aluko RE, Dhewa T, Moreno-Rojas JM. Potential Health Benefits of Plant Food-Derived Bioactive Components: An Overview. *Foods* 2021;10(4). doi: 10.3390/foods10040839.
49. Pandey KB, Rizvi SI. Plant polyphenols as dietary antioxidants in human health and disease. *Oxid Med Cell Longev* 2009;2(5):270-8. doi: 10.4161/oxim.2.5.9498.
50. Liu RH. Health-promoting components of fruits and vegetables in the diet. *Adv Nutr* 2013;4(3):384S-92S. doi: 10.3945/an.112.003517.
51. Nemeč MJ, Kim H, Marcianti AB, Barnes RC, Hendrick ED, Bisson WH, Talcott ST, Mertens-Talcott SU. Polyphenolics from mango (*Mangifera indica* L.) suppress breast cancer ductal carcinoma in situ proliferation through activation of AMPK pathway and suppression of mTOR in athymic nude mice. *J Nutr Biochem* 2017;41:12-9. doi: 10.1016/j.jnutbio.2016.11.005.
52. Moore MA, Park CB, Tsuda H. Soluble and insoluble fiber influences on cancer development. *Crit Rev Oncol Hematol* 1998;27(3):229-42. doi: 10.1016/s1040-8428(98)00006-7.
53. Buck K, Zaineddin AK, Vrieling A, Linseisen J, Chang-Claude J. Meta-analyses of lignans and enterolignans in relation to breast cancer risk. *Am J Clin Nutr* 2010;92(1):141-53. doi: 10.3945/ajcn.2009.28573.
54. Velentzis LS, Cantwell MM, Cardwell C, Keshtgar MR, Leathem AJ, Woodside JV. Lignans and breast cancer risk in pre- and post-menopausal women: meta-analyses of observational studies. *Br J Cancer* 2009;100(9):1492-8. doi: 10.1038/sj.bjc.6605003.
55. Touillaud MS, Thiebaut AC, Fournier A, Niravong M, Boutron-Ruault MC, Clavel-Chapelon F. Dietary lignan intake and postmenopausal breast cancer risk by estrogen and progesterone receptor status. *J Natl Cancer Inst* 2007;99(6):475-86. doi: 10.1093/jnci/djk096.

56. Rose DP, Vona-Davis L. The cellular and molecular mechanisms by which insulin influences breast cancer risk and progression. *Endocr Relat Cancer* 2012;19(6):R225-41. doi: 10.1530/ERC-12-0203.
57. Bowers LW, Rossi EL, O'Flanagan CH, deGraffenried LA, Hursting SD. The Role of the Insulin/IGF System in Cancer: Lessons Learned from Clinical Trials and the Energy Balance-Cancer Link. *Front Endocrinol (Lausanne)* 2015;6:77. doi: 10.3389/fendo.2015.00077.
58. Chajes V, Thiebaut AC, Rotival M, Gauthier E, Maillard V, Boutron-Ruault MC, Joulin V, Lenoir GM, Clavel-Chapelon F. Association between serum trans-monounsaturated fatty acids and breast cancer risk in the E3N-EPIC Study. *Am J Epidemiol* 2008;167(11):1312-20. doi: 10.1093/aje/kwn069.
59. Micha R, Mozaffarian D. Trans fatty acids: effects on metabolic syndrome, heart disease and diabetes. *Nat Rev Endocrinol* 2009;5(6):335-44. doi: 10.1038/nrendo.2009.79.
60. Matta M, Huybrechts I, Biessy C, Casagrande C, Yammine S, Fournier A, Olsen KS, Lukic M, Gram IT, Ardanaz E, et al. Dietary intake of trans fatty acids and breast cancer risk in 9 European countries. *BMC Med* 2021;19(1):81. doi: 10.1186/s12916-021-01952-3.
61. Rose DP. Effects of dietary fatty acids on breast and prostate cancers: evidence from in vitro experiments and animal studies. *Am J Clin Nutr* 1997;66(6 Suppl):1513S-22S. doi: 10.1093/ajcn/66.6.1513S.
62. Yland JJ, Wesselink AK, Lash TL, Fox MP. Misconceptions About the Direction of Bias From Nondifferential Misclassification. *Am J Epidemiol* 2022;191(8):1485-95. doi: 10.1093/aje/kwac035.

Table 1. Baseline characteristics of the study population overall and according to quintile of the healthful plant-based diet index, E3N cohort (N=65,574).<sup>1</sup>

	Quintile of the healthful plant-based diet index					
	Overall (N=65,574)	Q1 (N=12,822)	Q2 (N=12,832)	Q3 (N=14,577)	Q4 (N=13,008)	Q5 (N= 12,335)
Age, years	52.85 (6.65)	51.90 (6.57)	52.57 (6.64)	52.97 (6.64)	53.36 (6.68)	53.46 (6.57)
Educational level, n (%)						
Undergraduate or less	7,358 (11.22)	1,528 (11.91)	1,504 (11.72)	1,648 (11.31)	1,376 (10.58)	1,302 (10.56)
Graduate	34,927 (53.26)	6,926 (54.02)	6,829 (53.22)	7,768 (53.28)	7,006 (53.86)	6,398 (51.86)
Postgraduate	23,289 (35.52)	4,368 (34.07)	4,499 (35.06)	5,161 (35.41)	4,626 (35.56)	4,635 (37.58)
Alcohol intake, grams/day	11.58 (13.90)	13.17 (14.89)	12.48 (14.42)	11.57 (14.00)	10.81 (13.15)	9.80 (12.59)
Smoking status, n (%)						
Current	8,826 (13.46)	1,769 (13.80)	1,755 (13.68)	1,923 (13.19)	1,733 (13.32)	1,646 (13.35)
Former	21,348 (32.56)	3,846 (30.00)	4,019 (31.32)	4,786 (32.83)	4,377 (33.65)	4,320 (35.02)
Non-smoker	35,400 (53.98)	7,207 (56.20)	7,058 (55.00)	7,868 (53.98)	6,898 (53.03)	6,369 (51.63)
BMI, kg/m <sup>2</sup>	22.92 (3.22)	23.15 (3.48)	23.00 (3.31)	22.92 (3.14)	22.89 (3.13)	22.63 (3.00)
BMI categories, kg/m <sup>2</sup> , n (%)						
<20	8,826 (13.46)	1,769 (13.80)	1,755 (13.68)	1,923 (13.19)	1,733 (13.32)	1,646 (13.35)
20-24.99	21,348 (32.56)	3,846 (30.00)	4,019 (31.32)	4,786 (32.83)	4,377 (33.65)	4,320 (35.02)
≥25	35,400 (53.98)	7,207 (56.20)	7,058 (55.00)	7,868 (53.98)	6,898 (53.03)	6,369 (51.63)
Physical activity, MET-	49.24 (49.52)	47.66 (45.21)	48.55 (48.88)	48.90 (50.07)	49.65 (49.00)	51.57 (54.08)

h/week						
Energy intake (excluding alcohol) kcal/day	2129.36 (543.81)	2441.27 (543.17)	2226.76 (530.40)	2097.45 (510.86)	2001.04 (493.34)	1876.83 (463.09)
Age at menarche, years	12.78 (1.42)	12.83 (1.42)	12.80 (1.41)	12.79 (1.42)	12.76 (1.39)	12.73 (1.42)
Age at menopause, years	50.62 (3.82)	50.62 (3.82)	50.65 (3.81)	50.65 (3.81)	50.57 (3.82)	50.58 (3.84)
Age at first birth, n (%)						
< 30 years	51,384 (78.36)	10,237 (79.84)	10,182 (79.35)	11,345 (77.82)	10,099 (77.64)	9,521 (77.19)
≥ 30 years	6,625 (10.10)	1,365 (10.65)	1,262 (9.83)	1,527 (10.48)	1,297 (9.97)	1,174 (9.51)
Nulliparous	7,565 (11.54)	1,220 (9.51)	1,388 (10.82)	1,705 (11.70)	1,612 (12.39)	1,640 (13.30)
Breastfeeding, n (%)						
Ever	37,772 (57.60)	7,482 (58.35)	7,386 (57.56)	8,350 (57.28)	7,390 (56.81)	7,164 (58.08)
Never	24,380 (37.18)	4,682 (36.52)	4,785 (37.29)	5,475 (37.56)	4,916 (37.79)	4,522 (36.66)
Unknown	3,422 (5.22)	658 (5.13)	661 (5.15)	752 (5.16)	702 (5.40)	649 (5.26)
Ever use of menopausal hormone therapy, n (%)	19,761 (30.14)	3,404 (26.55)	3,812 (29.71)	4,417 (30.30)	4,203 (32.31)	3,925 (31.82)
Ever use of the contraceptive pill, n (%)	39,816 (60.72)	8,263 (64.44)	7,927 (61.78)	8,818 (60.49)	7,674 (58.99)	7,134 (57.84)
Past history of benign breast disease, n (%)	19,048 (29.05)	3,651 (28.47)	3,721 (29.00)	4,183 (28.70)	3,809 (29.28)	3,684 (29.87)
Family history of breast cancer, n (%)	4,841 (7.38)	931 (7.26)	929 (7.24)	1,135 (7.79)	952 (7.32)	894 (7.25)
Mammography in the last follow-up cycle,	44,751 (68.25)	8,512 (66.39)	8,766 (68.31)	9,910 (67.98)	9,041 (69.50)	8,522 (69.09)

n (%)						
Healthful plant-based diet index (continuous)	55.94 (6.07)	47.35 (2.82)	52.59 (1.02)	56.01 (0.97)	59.40 (1.01)	64.64 (2.78)

<sup>1</sup>Continuous variables were described using means and standard deviations, and categorical variables were described as numbers and percentages.

Table 2. Association of the healthful and unhealthful plant-based diet indices with overall breast cancer risk, E3N cohort (N=65,574).

			Model 1	Model 2	Model 3
	Number non-cases	Number cases	HR (95% CI)	HR (95% CI)	HR (95% CI)
<b>Healthful plant-based diet index</b>					
	N=61606	N=3968			
<i>P-non-linear</i>					< 0.01
Q1	11960	862	Reference	Reference	Reference
Q2	12002	830	0.93 (0.85, 1.03)	0.93 (0.84, 1.02)	0.93 (0.85, 1.03)
Q3	13760	817	0.78 (0.71, 0.86)	0.77 (0.70, 0.85)	0.79 (0.71, 0.87)
Q4	12290	718	0.77 (0.70, 0.85)	0.76 (0.69, 0.84)	0.78 (0.70, 0.86)
Q5	11594	741	0.84 (0.76, 0.93)	0.83 (0.75, 0.91)	0.86 (0.77, 0.95)
<i>P-trend</i>			< 0.01	< 0.01	< 0.01
<b>Unhealthful plant-based diet index</b>					
<i>P-non-linear</i>					0.18
1-SD increase	N=61606	N=3968	1.01 (0.98, 1.04)	1.02 (0.98, 1.05)	1.04 (1.01, 1.08)
Q1	11547	753	Reference	Reference	Reference
Q2	12955	813	0.96 (0.87, 1.06)	0.96 (0.87, 1.06)	0.98 (0.89, 1.09)
Q3	11575	711	0.94 (0.85, 1.04)	0.95 (0.85, 1.05)	0.99 (0.89, 1.10)
Q4	13414	837	0.96 (0.87, 1.05)	0.96 (0.87, 1.06)	1.02 (0.92, 1.13)
Q5	12115	854	1.10 (1.00, 1.22)	1.11 (1.01, 1.23)	1.20 (1.08, 1.33)
<i>P-trend</i>			0.12	0.06	<0.01

M1: Adjusted for age (as the time scale), stratified by birth cohort

M2: M1+ educational level, physical activity, smoking status, family history of breast cancer, breastfeeding, age at menarche, age at first full-term birth, past history of benign breast disease, ever use of the contraceptive pill, ever use of menopausal hormone therapy, and mammography in the last follow-up cycle

M3: M2 + body mass index, energy intake, and alcohol

Figure 1. Associations of the healthful plant-based diet index with BC fitted with restricted cubic splines (N=65,574; 5 knots placed at the 5th, 27.5th, 50th, 72.5th, and 95<sup>th</sup> percentiles),  $P_{\text{non-linear}} < 0.01$ . Risk estimates were adjusted for age (as the time scale), educational level, physical activity, smoking status, family history of BC, breastfeeding, age at menarche, age at first full-term birth, past history of benign breast disease, ever use of the contraceptive pill, ever use of menopausal hormone therapy, mammography in the last follow-up cycle, body mass index, energy intake, and alcohol (model stratified by birth cohort). The solid line represents the hazard ratio, and the dashed lines the lower and upper 95% confidence interval.

Figure 2. Associations of the unhealthful plant-based diet index with BC fitted with restricted cubic splines (N=65,574; 5 knots placed at the 5th, 27.5th, 50th, 72.5th, and 95<sup>th</sup> percentiles),  $P_{\text{non-linear}} = 0.18$ . Risk estimates were adjusted for age (as the time scale), educational level, physical activity, smoking status, family history of BC, breastfeeding, age at menarche, age at first full-term birth, past history of benign breast disease, ever use of the contraceptive pill, ever use of menopausal hormone therapy, mammography in the last follow-up cycle, body mass index, energy intake, and alcohol (model stratified by birth cohort). The solid line represents the hazard ratio, and the dashed lines the lower and upper 95% confidence interval.

Figure 3. The healthful and unhealthful plant-based diet indices and breast cancer risk, overall and subtypes, E3N cohort (N=65,574). Hazard ratios (Model 3)<sup>1</sup> for the highest (Q5) vs. the lowest (Q1) quintiles was presented in the figure. HR, hazard ratio, CI, confidence interval, ER, estrogen receptor, PR, progesterone receptor.

<sup>1</sup>HR adjusted for age (as the time scale), educational level, physical activity, smoking status, family history of breast cancer, breastfeeding, age at menarche, age at first full-term birth, past history of benign breast disease, ever use of the contraceptive pill, ever use of

menopausal hormone therapy, mammography in last follow-up cycle, body mass index, energy intake, and alcohol (model stratified by birth cohort).





