Insulinemic potential of diet and risk of total and subtypes of breast cancer among US females

Andrea Romanos-Nanclares,[1] *Fred K Tabung,*[2,3,4] *Walter C Willett,*[1,2,5] *Bernard Rosner,*[1,6] *Michelle D Holmes,*[1,5] *Wendy Y Chen,*[1,7] *Rulla M Tamimi,*[8] *and A Heather Eliassen*[1,2,5]

[1]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA; [2]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [3]Division of Medical Oncology, Department of Internal Medicine, The Ohio State University College of Medicine, Columbus, OH, USA; [4]The Ohio State University Comprehensive Cancer Center—Arthur G. James Cancer Hospital and Richard J. Solove Research Institute, Columbus, OH, USA; [5]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [6]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [7]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA; and [8]Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, USA

Abstract

Background: Insulin resistance and hyperinsulinemia play important roles in the progression of multiple chronic disease and conditions. Diet modulates insulin response; however, evidence is limited regarding whether diets with higher insulinemic potential increase the risk of invasive breast cancer.

Objectives: We aimed to prospectively evaluate the association between a food-based empirical dietary index for hyperinsulinemia (EDIH) and the incidence of invasive breast cancer.

Methods: We prospectively followed 76,686 women from the Nurses' Health Study (NHS; 1984–2016) and 93,287 women from the Nurses' Health Study II (NHSII; 1991–2017). Diet was assessed by food-frequency questionnaires every 4 y. The insulinemic potential of diet was evaluated using the previously established EDIH based on circulating C-peptide concentrations. Higher scores indicate higher insulinemic potential of the diet. Covariates included reproductive, hormonal, and anthropometric factors (height and BMI at age 18 y); race; socioeconomic status; total alcohol intake; total caloric intake; and physical activity.

Results: During 4,216,106 person-years of follow-up, we documented 10,602 breast cancer cases (6689 NHS, 3913 NHSII). In the pooled multivariable-adjusted analyses, women in the highest, compared with the lowest, EDIH quintile (Q) were at higher breast cancer risk ($HR_{Q5\ vs.\ Q1} = 1.15$; 95% CI: 1.07, 1.24; P-trend < 0.01). Although heterogeneity by estrogen receptor (ER) status was nonsignificant, the strongest association between EDIH and breast cancer was observed for ER-negative tumors ($HR_{Q5\ vs.\ Q1} = 1.21$; 95% CI: 1.00, 1.46; P-trend $= 0.02$). Among tumor molecular subtypes, the strongest associations were observed for human epidermal growth factor receptor 2 (HER2)–enriched tumors ($HR_{Q5\ vs.\ Q1} = 1.62$; 95% CI: 1.01, 2.61; P-trend $= 0.02$).

Conclusions: A dietary pattern contributing to hyperinsulinemia and insulin resistance was associated with greater breast cancer risk, especially ER-negative and HER2-enriched tumors. Our findings suggest that dietary modifications to reduce insulinemic potential may reduce the risk of breast cancer. *Am J Clin Nutr* 2022;116:1530–1539.

Keywords: breast cancer, epidemiology, diet, dietary patterns, hyperinsulinemia, cohort study

Introduction

Poor dietary quality has been linked to the development of many chronic diseases including breast cancer; hyperinsulinemia and insulin resistance may be important underlying mechanisms for this association (1–3). Although certain dietary factors appear to influence insulin resistance and secretion (4, 5), dietary patterns or indices that incorporate and account for the complex

Supported by grants UM1 CA186107, U01 CA176726, P01 CA87969, and R01 CA50385 from the National Institutes of Health, the Breast Cancer Research Foundation, Ramon Areces Foundation and Susan G Komen Foundation. The funding sources did not participate in the design or conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Supplemental Figure 1 and Supplemental Tables 1–6 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Address correspondence to AR-N (e-mail: nharo@channing.harvard.edu).

Abbreviations used: AHEI, Alternate Healthy Eating Index; AMED, Alternate Mediterranean Diet Score; CK, cytokeratin; DASH, Dietary Approaches to Stop Hypertension; EDIH, empirical dietary index for hyperinsulinemia; EDIP, Empirical Dietary Inflammatory Pattern; ER, estrogen receptor; FFQ, food-frequency questionnaire; HER2, human epidermal growth factor receptor 2; IGF-I, insulin-like growth factor I; IR, insulin receptor; MAPK, mitogen-activated protein kinase; NHS, Nurses' Health Study; NHSII, Nurses' Health Study II; PR, progesterone receptor; Q, quintile; TMA, tumor microarray.

Received July 1, 2022. Accepted for publication September 28, 2022.

First published online September 30, 2022; doi: https://doi.org/10.1093/ajcn/nqac284.

interactions among diverse components may be more predictive of diet–disease associations (6).

Several breast cancer risk factors, such as obesity and physical inactivity, contribute to insulin resistance, which, in turn, increases insulinemia (7, 8). Diets modulating these biological pathways of insulin response may influence breast cancer incidence. In epidemiologic studies, healthy dietary patterns, such as the Alternate Healthy Eating Index (AHEI), Alternate Mediterranean Diet (aMED), or Dietary Approaches to Stop Hypertension (DASH), among others, have been inversely associated with the risk of breast cancer (9–12). Also, greater adherence to a prudent dietary pattern, composed of vegetables, fruits, legumes, and whole grains, and lower adherence to a Western dietary pattern, high in saturated fats and red and processed meats as well as added sugars, fried foods, and refined grains, were associated with a reduced risk of breast cancer (13). Although these dietary patterns are associated with insulin response (14–16) they may not comprehensively capture the insulinemic potential of the diet, which may be an important mechanism linking dietary quality and breast cancer.

Biomarker-based dietary patterns may be crucial in this regard. Recently, we developed and evaluated the validity of an empirical index to assess the insulinemic potential of a usual diet, the empirical dietary index for hyperinsulinemia (EDIH), based on fasting C-peptide concentrations (17). C-peptide is co-secreted with insulin but has a longer half-life in plasma than insulin; consequently, it can serve as a marker to assess a time-integrated measure of insulinemia due to both the dietary glycemic load and insulin resistance. The glycemic index and insulin index assess the acute postprandial glycemic or insulinemic potential of specific foods (17) and have not predicted fasting C-peptide concentrations. C-peptide concentrations were positively associated with breast cancer in a recent meta-analysis (18). In the Nurses' Health Studies, higher C-peptide concentrations were associated with higher breast cancer risk in pre- and postmenopausal women, and this association was stronger for estrogen receptor (ER)–negative tumors (19). Serum C-peptide concentration may be modified by diet, and a dietary pattern consisting of a higher consumption of animal-based foods and refined grains has been associated with higher C-peptide concentrations (20, 21), whereas whole grains (5), fruits, vegetables (22), and coffee (4,20) have generally been associated with lower concentrations of insulin or C-peptide. A dietary pattern may be more appropriate than evaluating single nutrients or foods, which separately may not affect insulin concentrations sufficiently to influence breast cancer appreciably.

Although the EDIH has been associated with the risk of several major chronic diseases (23–31), its relation to breast cancer has not yet been evaluated. Given the potential role of hyperinsulinemia on breast cancer risk, we hypothesized that a greater insulinemic potential of diet, represented by higher EDIH scores, would be associated with a higher risk of breast cancer.

Methods

Study design

The Nurses' Health Study (NHS) is an ongoing study of 121,701 female nurses aged 30–55 y in 1976, and the Nurses' Health Study II (NHSII) has followed 116,429 female nurses (aged 25–42 y) since 1989. Every 2 y, participants have provided information on health-related factors and medical history. Women were followed from 1984 to 2016 in the NHS and from 1991 to 2017 in the NHSII. In this analysis, we excluded participants who had missing values in the dietary score (i.e., EDIH), and who had cancer (except for nonmelanoma skin cancer) or implausible energy intake (<600 or >3500 kcal/d) at baseline. Overall, a total of 76,686 women from the NHS and 93,287 from NHSII were included in the analysis (**Supplemental Figure 1**). The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and those of participating registries as required.

Diet assessment

Diet was assessed with semi-quantitative food-frequency questionnaires (FFQs) administered in the NHS in 1984, 1986, and every 4 y thereafter, and in the NHSII in 1991 and every 4 y thereafter. The numbers of FFQ food items have evolved: in the NHS, there were 116 items in 1984 and 1986 and ≥130 items thereafter; in the NHSII, the FFQ from 1991 had ≥130 items. The FFQs included foods with a specified standard portion size, and participants were asked to indicate the average consumption of each food during the previous year (from among 9 choices ranging from "almost never" to ">6/day").

The EDIH was constructed using weighted sums of food groups that predicted plasma C-peptide from a sample of 5812 women in NHS (17) and validated using fasting plasma C-peptide samples in the NHSII and the Health Professionals Follow-Up Study (HPFS) (17). Briefly, the EDIH was derived based on 39 predefined food groups from FFQs using stepwise regression models to identify a dietary pattern most predictive of circulating C-peptide concentrations beyond the postprandial state and which captures factors that also influence insulin resistance (e.g., coffee does not contribute calories or carbohydrates but can improve insulin sensitivity). Therefore, the EDIH may more exhaustively capture the influence of the whole diet on insulin secretion and blood concentrations. In total, 18 food groups (**Supplemental Table 1**) were identified that contribute either positively to the EDIH, including red meat, cream soups, margarine, butter, processed meat, and fruit juice, or inversely to the EDIH, such as whole grains, coffee, wine, and green-leafy vegetables. The EDIH showed low-to-moderate correlations with conventional dietary pattern scores (AHEI, aMED, and DASH; $r = -0.09$ to -0.45), and a strong inverse correlation with the healthful plant-based diet index ($r = -0.58$). Also, the EDIH has a positive correlation with the Western dietary pattern ($r = 0.63$) (25). In the present study, we calculated the EDIH score for each participant using FFQ data in each 4-y cycle.

Breast cancer ascertainment

We first identified incident breast cancer cases from biennial questionnaires. We requested permission from women reporting breast cancer to review hospital records and pathology reports for diagnosis confirmation and ascertainment of invasive vs. in situ and ER, progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status. Given the high confirmation rate of reported breast cancer cases in the NHS and NHSII

(>99%), we included both breast cancer cases confirmed via medical record review (>90%) and self-reported cases confirmed by the nurse but lacking a medical record. For deceased cases, the next of kin was contacted for permission. Deaths were reported by family members or by the postal service in response to follow-up questionnaires, or they were identified through the National Death Index.

Details of breast cancer tissue block collection and tumor microarray (TMA) construction have been reported previously (32). Briefly, we collected archived formalin-fixed, paraffin-embedded breast cancer blocks from participants with incident breast cancer diagnosed up through 2006. For molecular subtype classification, immunohistochemical staining information was available for the markers of ER, PR, HER2, cytokeratins (CK) 5/6, and epidermal growth factor receptor (EGFR). Further staining for the proliferative marker Ki-67 was achieved in NHS cases; Ki-67 data were not available for NHSII cases. Cases with TMAs were very similar to all suitable invasive cases in terms of demographics, breast cancer risk factors, and tumor characteristics.

Definitions that correlated with gene expression profile classifications were used for tumor molecular subtyping for a subgroup of cases. If Ki-67 expression data were missing (NHSII tumors), histological grade was used instead. Hence, luminal A tumors were ER-positive and/or PR-positive, HER2-negative, and Ki-67–negative (or histologic grade 1 or 2). Luminal B tumors were either *1*) ER-positive and/or PR-positive and HER2-positive or *2*) ER-positive and/or PR-positive, HER2-negative, and Ki-67–positive (or histologic grade 3). HER2-enriched tumors were ER-negative, PR-negative, and HER2-positive. Basal-like tumors were ER-negative, PR-negative, HER2-negative, and CK 5/6–positive and/or EGFR-positive. For evaluating ER-positive vs. ER-negative tumors, ER status was determined primarily from TMA slides and, if not available, secondarily from pathology reports. For cases diagnosed between 1980 and 2006 and in the NHS only, information was available for the insulin receptor (IR).

Ascertainment of covariates

Information on lifestyle and other potential risk factors was collected at baseline and updated biennially during follow-up through self-administered questionnaires, including race, socioeconomic status, age at menarche, age at menopause, postmenopausal hormone use, oral contraceptive use, parity, age at first birth, breastfeeding history, height, total alcohol intake, total caloric intake, physical activity, BMI at age 18 y, and change in weight since age 18 y. BMI was calculated as weight in kilograms divided by the square of the height in meters.

Statistical analysis

We calculated person-years of follow-up from the date of the return of the baseline questionnaire until the date of breast cancer diagnosis, other cancers (excluding nonmelanoma skin cancers), or the end of follow-up (June 2016 for NHS, June 2017 for NHSII), whichever occurred first. To better represent long-term exposures and reduce within-person variation, cumulative averages of EDIH were computed from all previous questionnaires up to the start of each 2-y follow-up interval. Similarly, the cumulative average intake of other covariates, when appropriate, was created to best reflect long-term food intake and lifestyle, and to minimize within-person variation.

To maximize the statistical power, data from the NHS and NHSII were pooled. Age-adjusted and multivariable-adjusted HRs and 95% CIs were calculated using a time-varying Cox regression model. All analyses were stratified by cohort and calendar year and age in months was the underlying time scale, enabling the finest possible control of confounding for age and secular trends. In the multivariable-adjusted model, we included race, socioeconomic status, age at menarche, age at menopause, postmenopausal hormone use, oral contraceptive use, parity, age at first birth, breastfeeding history, height, BMI at age 18, and total caloric intake (for categorizations, see footnote in table 2). We did not include weight gain after age 18 in the primary analysis because this could be a mediator of an effect of an hyperinsulinemic diet (23). We conducted a Wald test for trend, using the median EDIH value of each quintile, modeled as a continuous variable.

We performed several sensitivity analyses to test the robustness of our findings. First, we additionally adjusted for the consumption of food groups (coffee, red and processed meat) contributing most to the EDIH. Second, to take advantage of repeated diet assessments in these cohorts and evaluate the latency between these indices and breast cancer incidence, we conducted separate Cox models at different lag periods, with risk of invasive and ER-negative breast cancer. In the simple updated model, EDIH scores reported on the most recent FFQ before each follow-up interval were used; in the latency models, we used EDIH scores reported at different latencies (i.e., 4–8, 8–12, 12–16, 16–20 y) before breast cancer diagnosis. In addition, we applied a damped exponential weighting function that incorporates latency of exposure profiles that can be implemented using standard Cox regression software (33). Third, we tested whether the EDIH and breast cancer association varied by selected traditional risk factors: current BMI, physical activity, total alcohol intake, and menopausal status at diagnosis. Tests for interaction were obtained using the Wald test of cross-product interaction terms between the EDIH, modeled as a continuous variable, and potential effect modifiers. Fourth, to examine the independent association of the EDIH with breast cancer, we conducted a model further mutually adjusting for the empirical dietary inflammatory pattern (EDIP) (34). Mediation analyses (35) were performed to assess the extent to which associations may be statistically accounted for by weight gain from age 18 y and type 2 diabetes and estimated the mediated proportion (36, 37). To evaluate whether associations differed by molecular subtype or ER status, we used the Lunn-McNeil approach to derive the *P* value for heterogeneity (38).

Statistical tests were 2-sided, with *P* values <0.05 indicating statistical significance. All analyses were performed using SAS for UNIX version 9.4 (SAS Institute).

Results

Our study included 76,686 women from the NHS and 93,287 women from the NHSII. Compared with participants with lower EDIH scores, those in the highest quintile tended to have higher BMI (kg/m^2) and weight change

TABLE 1 Age and age-adjusted baseline characteristics of participants according to quintiles of the EDIH score in the NHS (1984) and the NHSII (1991)[1]

	NHS			NHSII		
	Q1 (n = 15,337)	Q3 (n = 15,337)	Q5 (n = 15,337)	Q1 (n = 18,657)	Q3 (n = 18,657)	Q5 (n = 18,657)
Median (range) EDIH score	0.1 (0.1, 0.2)	0.5 (0.4, 0.5)	0.9 (0.8,1)	0.2 (0.1, 0.2)	0.5 (0.4, 0.5)	0.9 (0.8, 1)
Age, y	52.3 (6.9)	51 (7.1)	49.2 (7.1)	37.5 (4.5)	36.6 (4.7)	36 (4.7)
BMI, kg/m^2	23.7 (3.7)	25.1 (4.6)	26.5 (5.7)	23 (4)	24.3 (5)	26.3 (6.5)
Waist circumference, cm	76.1 (9.4)	79 (10.5)	82.3 (12.5)	75.2 (10.5)	78.3 (12.1)	82.7 (14.5)
BMI at age 18 y, kg/m^2	21.3 (2.8)	21.4 (2.9)	21.5 (3.2)	21 (3)	21.2 (3.3)	21.7 (3.8)
Weight change from age 18 y, kg	6.3 (9.5)	9.9 (10.8)	13.5 (12.9)	5.4 (9.2)	8.5 (10.8)	12.7 (13.6)
Height, m	1.6 (0.1)	1.6 (0.1)	1.6 (0.1)	1.6 (0.1)	1.6 (0.1)	1.6 (0.1)
White, %	97.3	97.2	97.0	95.1	95.4	95.1
Self-reported history of diabetes, %	1.7	3.0	5.5	0.6	0.8	1.8
Family history of breast cancer, %	8.2	8.1	7.8	6.2	5.8	5.9
Biopsy-confirmed benign breast disease, %	28.0	25.8	24.7	9.9	9.4	9.4
Age at menarche <12 y, %	22.1	22.6	23.0	24.2	24.1	25.4
Oral contraceptive use, ever, %	49.5	49.5	49.6	84.0	85.0	84.5
Parous, %	91.6	93.0	93.1	64.7	76.8	78.1
Parity,[2] n[1]	3.1 (1.4)	3.2 (1.5)	3.2 (1.6)	2 (0.9)	2.1 (0.9)	2.2 (0.9)
Breastfeeding ≤6 mo,[2] %	36.3	36.8	36.3	65.9	62.9	58.0
Postmenopausal, %	48.6	48.4	48.4	2.9	3.3	3.4
Postmenopausal hormone use, never,[3] %	52.0	53.4	52.6	7.7	7.6	7.4
Physical activity, MET-h/wk	13.8 (14.4)	12 (12)	10.8 (12.1)	26.1 (32.8)	20.1 (24)	18.1 (23)
Dietary intake						
Total energy, kcal/d	1436 (457)	1683 (436)	2214 (504)	1459 (462)	1730 (446)	2278 (522)
Total carbohydrates, %	48.3 (8.9)	46.5 (7.3)	44.2 (7.8)	52.8 (8.5)	49.5 (6.7)	47.4 (7.6)
Total protein, %	16.8 (3.1)	18.1 (3.3)	18.3 (3.6)	18.4 (3.6)	19.7 (3.4)	19.4 (3.6)
Saturated fat, %	11.7 (3)	12.4 (2.4)	13.4 (2.4)	10.4 (2.9)	11.2 (2.2)	12 (2.2)
Monounsaturated fat, %	11.4 (2.4)	12.7 (2.2)	13.8 (2.3)	10.8 (2.6)	11.9 (2.2)	13.2 (2.3)
Polyunsaturated fat, %	6.5 (1.9)	6.7 (1.7)	6.7 (1.6)	5.5 (1.5)	5.6 (1.3)	5.8 (1.3)
Alcohol, g/d	10.3 (13.8)	6.0 (9)	5.9 (10.9)	4.7 (7)	2.7 (5.3)	2.5 (5.5)

[1]All variables are standardized to the age distribution of the study population, except for age. Values are means (SDs) for continuous variables and percentage of participants for categorical variables unless otherwise indicated. EDIH, empirical dietary index for hyperinsulinemia; MET, metabolic equivalent of task; NHS, Nurses' Health Study; NHSII; Nurses' Health Study II; Q, quintile. [2]Among parous women only. [3]Among postmenopausal.

since age 18 (kilograms), and lower physical activity levels, and were more likely to have a history of diabetes (**Table 1**). They also reported higher total caloric intake (kilocalories/day).

During 4,216,106 person-years of follow-up, we documented 10,602 breast cancer cases (6689 NHS, 3913 NHSII). Although the age-adjusted model showed no association, in the multivariable-adjusted analysis (**Table 2**), the EDIH was significantly associated with breast cancer risk [$HR_{Quintile\ (Q)\ 5\ vs\ .Q1}$ = 1.15; 95% CI: 1.07, 1.24; P-trend < 0.01]. Reproductive and hormonal factors were the main covariates responsible for the observed changes in the effect estimates between the age-adjusted and multivariable-adjusted models. Additional adjustment for weight change since age 18, a potential mediator of an effect of the EDIH on risk of breast cancer, and the EDIP modestly attenuated the magnitude of the association ($HR_{Q5vs.Q1}$ = 1.08; 95% CI: 1.00, 1.16; P-trend = 0.03; and $HR_{Q5vs.Q1}$ = 1.09; 95% CI: 1.00, 1.19; P-trend = 0.04, respectively) (data not shown). In a separate analysis by cohort (**Supplemental Table 2**), higher EDIH scores were associated with a higher risk in the NHS ($HR_{Q5vs.Q1}$ = 1.17; 95% CI: 1.07, 1.28; P-trend < 0.0001) and a suggestively higher risk in the NHSII ($HR_{Q5vs.Q1}$ = 1.11; 95% CI: 0.99, 1.25; P-trend = 0.10). In ancillary analyses, we took alcohol components out of the EDIH score and consequently adjusted for cumulatively updated total alcohol consumption in the multivariable model, but results were basically unchanged. Furthermore, we evaluated the extent to which the direct association with higher EDIH may be mediated by weight gain from age 18 or type 2 diabetes (**Supplemental Table 3**). The calculated mediation proportion was 37.4% (95% CI: 19%, 60.4%; P < 0.001) and 7.4% (95% CI: 3.5%, 14.8%; P < 0.001), indicating that weight gain and diabetes could statistically explain approximately 37% and 7% of the positive association with EDIH (Supplemental Table 3).

We observed no statistical evidence of interaction by BMI, waist circumference, physical activity, or alcohol intake (P-interaction > 0.05; **Supplemental Table 4**); nonetheless, the association between EDIH and breast cancer risk was suggestively stronger for postmenopausal women ($HR_{Q5vs.Q1}$ = 1.19; 95% CI: 1.10, 1.29; P-trend < 0.001, P-interaction = 0.05) (**Table 3**). The latter association remained significant when we

TABLE 2 Multivariable HRs (95% CIs) of overall and subtypes of breast cancer according to EDIH quintiles in the NHS and NHSII

	Quintiles of EDIH					*P*-trend[2]
	Q1	Q2	Q3	Q4	Q5	
Invasive breast cancer						
Cases/person-years	2259/843,605	2145/843,690	2081/843,656	2090/843,456	2027/841,701	
Age-adjusted	1.00	0.97 (0.91, 1.03)	0.96 (0.90, 1.01)	0.98 (0.92, 1.04)	0.97 (0.91, 1.03)	0.39
MV-adjusted	1.00	1.02 (0.96, 1.09)	1.04 (0.98, 1.11)	1.10 (1.03, 1.17)	1.15 (1.07, 1.24)	<0.01
ER-negative						
Cases/person-years	300/845,475	296/845,465	311/845,370	325/845,139	316/843,286	
Age-adjusted	1.00	0.99 (0.84, 1.16)	1.06 (0.90, 1.24)	1.12 (0.96, 1.32)	1.10 (0.94, 1.29)	0.11
MV-adjusted	1.00	1.01 (0.86, 1.19)	1.10 (0.93, 1.30)	1.19 (1.00, 1.42)	1.21 (1.00, 1.46)	0.02
ER-positive						
Cases/person-years	1570/844,297	1424/844,389	1409/844,323	1351/844,205	1312/842,383	
Age-adjusted	1.00	0.93 (0.86, 1.00)	0.93 (0.87, 1.00)	0.91 (0.84, 0.98)	0.91 (0.84, 0.98)	0.01
MV-adjusted	1.00	0.99 (0.92, 1.06)	1.03 (0.95, 1.11)	1.04 (0.96, 1.12)	1.11 (1.02, 1.22)	0.01
P-heterogeneity[3] = 0.30						
Luminal A[4]						
Cases/person-years	578/585,800	498/585,985	507/585,903	470/585,602	450/584,192	
Age-adjusted	1.00	0.89 (0.79, 1.00)	0.93 (0.82, 1.04)	0.88 (0.78, 1.00)	0.88 (0.78, 1.00)	0.24
MV-adjusted	1.00	0.93 (0.82, 1.05)	0.99 (0.87, 1.12)	0.97 (0.85, 1.11)	1.02 (0.88, 1.18)	0.06
Luminal B[4]						
Cases/person-years	250/586,097	228/586,205	222/586,160	211/585,833	214/584,405	
Age-adjusted	1.00	0.92 (0.77, 1.10)	0.91 (0.76, 1.09)	0.88 (0.73, 1.05)	0.91 (0.75, 1.09)	0.29
MV-adjusted	1.00	0.96 (0.80, 1.16)	0.99 (0.82, 1.19)	0.98 (0.80, 1.20)	1.07 (0.86, 1.33)	0.54
HER-2[4]						
Cases/person-years	43/586,282	45/586,380	43/586,322	57/585,958	53/584,553	
Age-adjusted	1.00	1.07 (0.70, 1.62)	1.07 (0.70, 1.64)	1.46 (0.98, 2.18)	1.43 (0.95, 2.15)	0.03
MV-adjusted	1.00	1.09 (0.71, 1.66)	1.12 (0.72, 1.75)	1.57 (1.02, 2.43)	1.62 (1.01, 2.61)	0.02
Basal-like[4]						
Cases/person-years	39/586,282	57/586,374	56/586,312	63/585,957	57/584,538	
Age-adjusted	1.00	1.48 (0.98, 2.23)	1.51 (1.00, 2.27)	1.74 (1.16, 2.60)	1.63 (1.08, 2.45)	0.02
MV-adjusted	1.00	1.40 (0.93, 2.12)	1.38 (0.90, 2.11)	1.58 (1.03, 2.43)	1.53 (0.96, 2.43)	0.10
P-heterogeneity[3] = 0.10						
Insulin receptor negative[5]						
Cases/person-years	179/298,160	192/298,147	182/298,044	210/297,786	157/297,046	
Age-adjusted	1.00	1.07 (0.87, 1.32)	1.06 (0.86, 1.30)	1.24 (1.02, 1.52)	0.97 (0.78, 1.21)	0.81
MV-adjusted	1.00	1.10 (0.89, 1.35)	1.09 (0.88, 1.35)	1.29 (1.03, 1.60)	1.02 (0.80, 1.31)	0.55
Insulin receptor positive[5]						
Cases/person-years						
Age-adjusted	1.00	0.91 (0.75, 1.10)	0.98 (0.81, 1.18)	0.83 (0.68, 1.02)	1.01 (0.83, 1.23)	0.89
MV-adjusted	1.00	0.94 (0.77, 1.14)	1.03 (0.84, 1.26)	0.90 (0.72, 1.12)	1.14 (0.91, 1.44)	0.34
P for heterogeneity = 0.80						

All analyses were conducted using Cox models stratified by cohort, age in months, and calendar year. The multivariable model adjusted for race (non-Hispanic White, Black, Asian-American, Hispanic White), age at menarche (<12, 12, 13, 14, >14 y), menopausal status, and age at menopause (premenopausal; <45, 45–49, 50–52, ≥53 y; unknown), postmenopausal hormone use [never user, past user, current user (estrogen only for <5 y), current user (estrogen only for ≥5 y), current estrogen + progestin user for <5 y, current estrogen + progestin user for ≥5 years, current user of other types], oral contraceptive use history (never, ever), parity and age at first birth (nulliparous, 1 child before age 25, 1 child at ≥25 years of age, ≥2 children before age 25, ≥2 children ≥25 y of age), breastfeeding history (never, breastfed for ≤6 mo, breastfed for >6 mo), family history of breast cancer (yes or no), history of biopsy-confirmed benign breast disease (yes or no), height (<1.60, 1.60–1.64, 1.65–1.69, 1.70–1.74, ≥1.75 m), cumulatively updated alcohol intake (0, <5, 5–9, 10–14, ≥15 g/d), cumulatively updated total caloric intake (kcal/d, quintiles), physical activity (continuous MET-hours/week), neighborhood-based socioeconomic status indicator (continuous), and BMI at age 18 (kg/m^2; <20.0, 20.0–21.9, 22.0–23.9, 24.0–26.9, ≥27.0). EDIH, empirical dietary index for hyperinsulinemia; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; MET, metabolic equivalent of task; MV, multivariable; NHS, Nurses' Health Study; NHSII; Nurses' Health Study II; Q, quintile.

[1]The *P*-trend was calculated by assigning the EDIH quintile medians to all the participants in the quintile and modeling as continuous variables.

[2]For testing heterogeneity by subtype, we used the Lunn–McNeil approach, for the multivariable model.

[3]Due to smaller sample sizes in analyses, to ensure that models would run, covariate categorizations were simplified.

[4]Insulin receptor (IR)–positive or -negative status was determined as ≥ (positive) or < (negative) the median of IR expression (cytoplasmic and membranous).

further adjusted for the EDIP (HR $_{Q5vs.Q1}$ = 1.13; 95% CI: 1.01, 1.25; *P*-interaction = 0.10) (data not shown).

In latency analyses (**Supplemental Table 5**) the highest compared with the lowest EDIH 0–16 y before diagnosis was associated with 15–23% higher risk of invasive breast cancer (HR: 1.15; 95% CI: 1.08, 1.24 for simple update or 0–4 y; HR: 1.15; 95% CI: 1.07, 1.23 for 4–8 y; HR: 1.23; 95% CI: 1.14, 1.33 for 8–12 y; and HR: 1.17; 95% CI: 1.07, 1.27 for 12–16 y).

TABLE 3 Multivariable HRs (95% CIs) of overall and ER-negative breast cancer risk according to quintiles of EDIH by menopausal status in the NHS and NHSII[1]

		Quintiles of EDIH					*P*-trend[2]	*P*-interaction[3]
	Cases, *n*	Q1	Q2	Q3	Q4	Q5		
Invasive breast cancer								
Menopausal status								
Premenopausal	2222	1.00	0.98 (0.86, 1.12)	1.04 (0.91, 1.19)	1.04 (0.90, 1.19)	1.07 (0.91, 1.25)	0.34	
Postmenopausal	7746	1.00	1.03 (0.96, 1.10)	1.03 (0.96, 1.11)	1.12 (1.04, 1.21)	1.19 (1.10, 1.29)	<0.001	0.05
ER-negative breast cancer								
Menopausal status								
Premenopausal	391	1.00	0.84 (0.60, 1.17)	1.00 (0.72, 1.39)	1.05 (0.75, 1.47)	1.01 (0.70, 1.47)	0.67	
Postmenopausal	1042	1.00	1.10 (0.90, 1.33)	1.09 (0.89, 1.33)	1.22 (0.99, 1.50)	1.29 (1.03, 1.62)	0.02	0.36

[1]The multivariable model adjusted for race (non-Hispanic White, Black, Asian-American, Hispanic White), age at menarche (<12, 12, 13, 14, >14 y), menopausal status, and age at menopause (premenopausal; <45, 45–49, 50–52, ≥53 y; unknown), postmenopausal hormone use [never user, past user, current user (estrogen only for <5 y), current user (estrogen only for ≥5 y), current estrogen + progestin user for <5 y, current estrogen + progestin user for ≥5 years, current user of other types], oral contraceptive use history (never, ever), parity and age at first birth (nulliparous, 1 child before age 25, 1 child at ≥25 years of age, ≥2 children before age 25, ≥2 children ≥25 y of age), breastfeeding history (never, breastfed for ≤6 mo, breastfed for >6 mo), family history of breast cancer (yes or no), history of biopsy-confirmed benign breast disease (yes or no), height (<1.60, 1.60–1.64, 1.65–1.69, 1.70–1.74, ≥1.75 m), cumulatively updated alcohol intake (0, <5, 5–9, 10–14, ≥15 g/d), cumulatively updated total caloric intake (kcal/d, quintiles), physical activity (continuous MET-hours/week), neighborhood-based socioeconomic status indicator (continuous), and BMI at age 18 (kg/m^2; <20.0, 20.0–21.9, 22.0–23.9, 24.0–26.9, ≥27.0). EDIH, empirical dietary index for hyperinsulinemia; ER, estrogen receptor; NHS, Nurses' Health Study; NHSII; Nurses' Health Study II; Q, quintile.

[2]The *P*-trend was calculated by assigning the median to all the participants in the quintile and modeling as continuous variables.

[3]*P*-interaction was calculated using the Wald test by including the interaction term.

In subgroup analyses of a factor, that potential effect modifier was not additionally adjusted for in the multivariable model.

Twenty-one MET-h/week is equivalent to ~7 h/wk of brisk walking.

We observed no significant heterogeneity by ER status (*P*-heterogeneity = 0.30) (Table 2), although direct associations were stronger for ER-negative ($HR_{Q5vs.Q1}$ = 1.21; 95% CI: 1.00, 1.46; *P*-trend = 0.02) than for ER-positive ($HR_{Q5vs.Q1}$ = 1.11; 95% CI: 1.02, 1.22; *P*-trend = 0.01) breast cancers. Further adjustment for weight change since age 18 did not materially alter the association for ER-negative breast cancer ($HR_{Q5vs.Q1}$ = 1.18; 95% CI: 0.97, 1.43; *P*-trend = 0.045), but it did change for ER-positive breast cancer ($HR_{Q5vs.Q1}$ = 1.04; 95% CI: 0.95, 1.13; *P*-trend = 0.39) (data not shown). Weight gain could statistically explain 59.2% (15.9%, 91.7%; $P < 0.001$) of the positive association for ER-positive tumors and might partly mediate that association (Supplemental Table 3). Although *P*-heterogeneity was 0.10 across the molecular subtypes (Table 2), significant associations were observed with HER2-enriched tumors ($HR_{Q5vs.Q1}$ = 1.62; 95% CI 1.01, 2.61; *P*-trend = 0.02). We did not observe any association between cumulative average EDIH and breast cancer based on IR status (*P*-heterogeneity = 0.80; Table 2).

Additional adjustment for the consumption of food groups (coffee, red and processed meat) contributing most to the EDIH and diabetes (**Supplemental Table 6**) or any other food components of the EDIH score (data not shown) did not alter the results. Furthermore, we adjusted for the EDIP to assess the independent contribution of hyperinsulinemia for ER-negative breast cancer, and the association was attenuated and no longer significant ($HR_{Q5vs.Q1}$ = 1.04; 95% CI: 0.82, 1.30; *P*-trend = 0.72) (data not shown). We also assessed the extent to which the association of ER-negative breast cancer with higher EDIH may be mediated by weight gain from age 18 y and diabetes. The proportion of EDIH association statistically accounted for by weight change since age 18 was 23.3% (95% CI: 4.6%, 65.7%; $P = 0.04$). In addition, the percentage of exposure effect was too small (<1%) to calculate reliably, and diabetes was not an intermediate to EDIH (Supplemental Table 3).

When we examined EDIH at specific lags (Supplemental Table 5), we identified a lag period for the association of EDIH and ER-negative breast cancer at various periods. The direct association between EDIH and ER-negative breast cancer was the strongest when EDIH was assessed 0–4 y (HR: 1.25; 95% CI: 1.05, 1.50), 4–8 y (HR: 1.34; 95% CI: 1.10, 1.62), and 8–12 y (HR: 1.31; 95% CI: 1.05, 1.63) before diagnosis. These results also suggested a 12–16-y lag for the association between EDIH and ER-negative breast cancer (HR: 1.18; 95% CI: 0.92, 1.52); however, the magnitude of association was less strong, and there was no association at later time points (16–20-y lag). Overall, the lag-specific results agree with the latency analyses (damped exponential weighting model), which suggested, although nonsignificantly, that recent adherence to the EDIH may be a more important predictor of risk than more distant exposure for ER-negative breast cancer (data not shown).

Discussion

In these 2 large prospective US cohorts, we investigated the associations of a dietary pattern reflecting the contribution of foods to hyperinsulinemia and insulin resistance with the risk of breast cancer. This association was most evident in relation to ER-negative and HER2-enriched tumors. To our knowledge, this is the first prospective cohort study examining dietary insulinemic potential and breast cancer development. The strong positive associations for ER-negative breast tumors remained significant after further adjusting for weight change or foods contributing most to the EDIH. Overall, these findings support the importance of the insulin signaling pathway for the etiology of breast cancer.

Dietary patterns with higher insulinemic potential are rich in red and processed meat, sugar-sweetened beverages, and French fries, but low in whole fruit, whole grains, high-fat dairy products (e.g., cheese, whole milk, yogurt), green-leafy vegetables, coffee, and wine. The resultant nutrient profile is rich in total and saturated fat, cholesterol, and total and animal protein and low in fiber. In the current analyses, positive associations were found for the EDIH with invasive breast cancer. Adiposity is closely related to diet and can, in part, mediate its role in breast cancer, with hyperinsulinemia/insulin resistance promoting breast cancer cell growth (39, 40). In previous studies, higher EDIH scores were associated with substantial long-term weight gain (23); therefore, we further adjusted for weight change to highlight the mediating influence of adiposity on the association between EDIH and breast cancer risk. As predicted, the influence of the dietary pattern was strongly mediated by weight change for invasive breast cancer. Indeed, increasing the likelihood of weight gain since age 18 y explained approximately 40% of the direct association with EDIH adherence. Higher diabetes prevalence statistically explained only 7.4% of the positive association with EDIH and we observed no interaction by diabetes status, which is consistent with our previous results (41). Furthermore, our results were slightly attenuated but remained significant after accounting for the inflammatory potential of the diet, which might reflect an independent association of insulinemic diets and overall invasive breast cancer risk.

Estrogen exposure is an established driver of breast cancer. Thus, any potential influence of dietary factors in ER-positive tumors may be hard to identify given the strong influence of hormonal factors. The association between EDIH and ER-positive tumors was attenuated after adjusting for weight change. On the other hand, in ER-negative tumors, other factors, such as diet, may exert a relatively greater influence and be more easily visible. In fact, the strong positive association between EDIH and ER-negative breast cancer remained after adjusting for weight change, suggesting that this empirical hypothesis-oriented dietary index for hyperinsulinemia has an impact independent of adiposity, possibly through mitogen-activated protein kinase (MAPK) (42) and PI3K/Akt/mTOR (43) pathways (44, 45). This is also consistent with a nested case-control study within the Nurses' Health Studies (19) in which C-peptide concentrations were more strongly associated with risk of ER-negative breast cancer than with ER-positive cases. Notably, the positive association of EDIH and subsequent ER-negative breast cancer was attenuated after accounting for the inflammatory potential of diet. Given these results, the insulinemic potential of diet might not play a unique and strong role in ER-negative breast cancer when accounting EDIP and therefore focusing on the independent information not available in the EDIP.

Our finding of stronger direct EDIH associations with ER-negative breast cancers is consistent with other studies that have reported lower risks of ER-negative breast cancer associated with lower glycemic load (46), lower total carbohydrate intake (46), higher dietary fiber (47), and higher adherence to a priori scores (DASH, Recommended Food Scores, aMED, healthful plant-based diet index) favoring plant foods (11). Nonetheless, the ability of these diet quality scores to predict breast cancer risk depends on how well these scores measure dietary risk factors for breast cancer. In this regard, the dietary score in our study includes several food groups that are not regularly captured in other dietary patterns, including certain vegetable categories (e.g., green-leafy vegetables), coffee, and animal-based foods besides red and processed meat, such as eggs. Some of these foods have been associated with risk of breast cancer, particularly for more aggressive tumors such as ER-negative, basal-like, or HER2-enriched breast tumors (11, 48, 49). In fact, our results are in close agreement with previous findings, which could be expected, as most healthy plant foods that are inversely weighted in the EDIH [e.g., green-leafy vegetables (48), whole fruits (48), coffee (41, 50)] have been associated with a lower risk of breast cancer in prospective cohort studies, including our own. Moreover, a healthful plant-based diet, inversely correlated with the EDIH ($r = –0.58$), and emphasizing the quality of plant foods (such as whole grains and foods rich in dietary fiber) was recently associated with a decreased risk of ER-negative breast tumors (51).

However, some components of the index are weighted towards components that are risk factors for breast cancer (52), such as alcohol consumption and, more specifically, red and white wine. Therefore, we also analyzed the EDIH without the alcohol component and controlled for alcohol as a confounder in the statistical model and observed similar results. Moreover, positively scored food groups such as red and processed meat, low-energy and high-energy beverages, butter, French fries, and cream soups have been associated with an increased risk of breast cancer (53–55). Thus, dietary patterns that do not account for these food items may be missing crucial elements of diet that influence breast cancer risk and could partly justify previous null and mixed findings.

Because ER-negative tumors tend to be more aggressive and have fewer treatment options than ER-positive tumors (45), it is of special importance to identify prevention strategies. Moreover, among the ER-negative breast cancers, a higher EDIH was associated with a higher risk for HER2-enriched tumors and a suggestively higher risk for basal-like tumors. Factors related to adiposity/hyperinsulinemia may be important for these subtypes (45), especially for triple-negative breast cancers, of which 80% are basal-like tumors. Previous studies examining associations between dietary patterns and molecular subtypes of breast cancer are limited. In the Nurses' Health Studies, and consistent with our current results, we observed an inverse trend for the DASH dietary pattern and HER2-enriched breast cancer (11).

Insulin resistance is tethered to hyperinsulinemia, and the direct mitogenic effect of insulin and the indirect effect through increased production of insulin-like growth factor I (IGF-I) and reduction of IGF-I binding proteins might help to explain the link between hyperinsulinemia and breast cancer (56). Free IGF-I has mitogenic and antiapoptotic effects and has been suggested to be associated with higher breast cancer risk (57). Higher adherence to an insulinemic diet may lead to increased risk of ER-negative breast cancer by potentiation of the IR (58). Insulin binding to IR leads to downstream activation of PI3K/AKt and MAPK signaling pathways (59), which induces mTOR signaling to promote cell growth (60). The latter can also stimulate mitochondrial biogenesis and activity, which, in turn, increases TCA cycle utilization and ATP production [through increased rates of oxidative phosphorylation (61)]. Since insulin can activate the PI3K/AKT/mTOR pathways, it is predicted that hyperinsulinemia, in the absence of obesity

and type 2 diabetes, may drive the aggressive biology of ER-negative tumors (62). Because we observed a stronger association with ER-negative than with ER-positive breast cancer, our results suggest that the putative causal effect of EDIH on breast cancer promotion may be exerted through non-estrogenic pathways. Identifying specific biological pathways through which diet and insulin may act to influence breast cancer development will further elucidate mechanisms of action to guide population-based policies towards effective prevention and intervention strategies. Integrating diet and metabolomics data may be helpful in this regard. For example, a low-quality dietary pattern (high EDIH) has been associated with high concentrations of different acylcarnitines (63), which may increase breast cancer risk (64). More specifically, His et al. (65) found that higher concentrations of acylcarnitine C2, which are a marker for lipid oversupply (that could enhance cancer cell proliferation) and upregulate fatty acid oxidation, were positively associated with breast cancer risk.

Limitations and strengths

This study's strengths include the 2 large prospective cohorts, detailed and updated dietary and covariate information, and availability of tissue information for the determination of molecular subtypes. The exposures were cumulatively updated over time and have been validated, minimizing misclassification, and reflecting long-term dietary patterns. The EDIH has previously been strongly associated with substantial long-term weight gain in men and women, independently of total energy intake (23), type 2 diabetes (27, 28), and several cancers (24–26, 29–31), indicating that the score is well measured.

However, our study has limitations. First, inevitable measurement errors in assessing diet, which would likely be nondifferential in relation to risk of breast cancer, may have caused underestimation of associations. Nonetheless, previous validation studies have shown reasonably good correlations between FFQs and diet records, suggesting that dietary intake is generally well measured (66, 67). Because the EDIH was empirically derived from C-peptide data, the strength of the association between the score and breast cancer relies not only on the association of index and biomarkers but also on the strength of association between C-peptide and breast cancer (18, 19). Nonetheless, EDIH scores are robustly associated with C-peptide in independent validation datasets (17). We additionally used multiple FFQs over follow-up to reduce within-person variation and better approximate habitual long-term diet. The EDIH scores were cumulatively averaged from multiple time points, which is likely more relevant to the natural course of breast cancer that spans several years. Another limitation was that we could not evaluate diet from childhood/adolescence, which may be a critical period for breast cancer initiation. Also, we had limited power to evaluate certain molecular subtypes. In addition, residual confounding cannot be excluded, but we controlled for a wide variety of breast cancer risk factors. Moreover, because the participants were mostly White, the results may not be generalizable to populations with different underlying breast cancer risks. Nonetheless, it is unlikely that the biological mechanisms would differ qualitatively in other populations, although the magnitudes of the association might vary (68).

Conclusions

In conclusion, in this large prospective study, higher hyperinsulinemic dietary scores (reflecting higher dietary insulinemic potential) were associated with a greater risk of invasive breast cancer, with a large part of this association being explained by weight gain. However, independent of adiposity, higher adherence to the EDIH was associated with higher ER-negative breast cancer incidence. Thus, dietary recommendations emphasizing the importance of avoiding high insulinemic dietary patterns (i.e., processed meat, red meat, high- and low-energy beverages) and prioritizing low insulinemic dietary patterns (i.e., whole fruit, green-leafy vegetables, full-fat dairy, and coffee) as one of the important components of a healthy diet could be considered for the primary prevention of breast cancer. Further studies may help identify specific biological pathways through which an insulinemic diet is implicated in breast cancer development.

We acknowledge the contribution to this study from central cancer registries supported through the CDC's National Program of Cancer Registries (NPCR) and/or the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries include the following: Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Maine, Maryland, Michigan, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, Seattle SEER Registry, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia, and Wyoming.

The authors' responsibilities were as follows—AR-N: conceptualization, resources, data curation, software, formal analysis, investigation, visualization, methodology, and writing (original draft, review and editing); FKT, WCW, BAR, and RMT: Methodology and writing (review and editing); WYC and MDH: writing (review and editing); AHE: conceptualization, resources, data curation, supervision, and writing (review and editing); all authors: assume full responsibility for analyses and interpretation of these data and read and approved the final manuscript. The authors report no conflicts of interest.

Data Availability

The data described in the article, code book, and analytic code will be made available upon application and approval. Further information including the procedures to obtain and access data from the Nurses' Health Studies is described at https://www.nurseshealthstudy.org/researchers (e-mail: nhsaccess@channing.harvard.edu).

References

1. Giovannucci E, Harlan DM, Archer MC, Bergenstal RM, Gapstur SM, Habel LA, et al. Diabetes and cancer: a consensus report. Diabetes Care 2010;33:1674–85.
2. Yee LD, Mortimer JE, Natarajan R, Dietze EC, Seewaldt VL. Metabolic health, insulin, and breast cancer: why oncologists should care about insulin. Front Endocrinol (Lausanne) 2020;11:58.
3. Hu Y, Zhang X, Ma Y, Yuan C, Wang M, Wu K, et al. Incident type 2 diabetes duration and cancer risk: a prospective study in two US cohorts. J Natl Cancer Inst 2021;113:381–9.
4. Wu T, Willett WC, Hankinson SE, Giovannucci E. Caffeinated coffee, decaffeinated coffee, and caffeine in relation to plasma C-peptide levels, a marker of insulin secretion, in U.S. women. Diabetes Care 2005;28:1390–6.

5. Jensen MK, Koh-Banerjee P, Franz M, Sampson L, GrãNbã¦K M, Rimm EB. Whole grains, bran, and germ in relation to homocysteine and markers of glycemic control, lipids, and inflammation. Am J Clin Nutr 2006;83:275–83.
6. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002;13:3–9.
7. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature 2006;444:840–6.
8. Hamburg NM, Mcmackin CJ, Huang AL, Shenouda SM, Widlansky ME, Schulz E, et al. Physical inactivity rapidly induces insulin resistance and microvascular dysfunction in healthy volunteers. Arterioscler Thromb Vasc Biol 2007;27:2650–6.
9. Petimar J, Park Y-MM, Smith-Warner SA, Fung TT, Sandler DP. Dietary index scores and invasive breast cancer risk among women with a family history of breast cancer. Am J Clin Nutr 2019;109:1393–401.
10. Fung TT, Hu FB, Mccullough ML, Newby PK, Willett WC, Holmes MD. Diet quality is associated with the risk of estrogen receptor–negative breast cancer in postmenopausal women. J Nutr 2006;136:466–72.
11. Hirko KA, Willett WC, Hankinson SE, Rosner BA, Beck AH, Tamimi RM, et al. Healthy dietary patterns and risk of breast cancer by molecular subtype. Breast Cancer Res Treat 2016;155:579–88.
12. Harris HR, Willett WC, Vaidya RL, Michels KB. Adolescent dietary patterns and premenopausal breast cancer incidence. Carcinogenesis 2016;37:376–84.
13. Xiao Y, Xia J, Li L, Ke Y, Cheng J, Xie Y, et al. Associations between dietary patterns and the risk of breast cancer: a systematic review and meta-analysis of observational studies. Breast Cancer Res 2019;29;21:16.
14. Fung TT. Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women. Arch Intern Med 2008;168:713–20.
15. Fung TT, Rimm EB, Spiegelman D, Rifai N, Tofler GH, Willett WC, et al. Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. Am J Clin Nutr 2001; 73:61–7.
16. Fung TT, McCullough ML, Newby PK, Manson JE,Meigs JB,Rifai N, et al. Diet-quality scores and plasma concentrations of markers of inflammation and endothelial dysfunction. Am J Clin Nutr 2005;82:163–73.
17. Tabung FK, Wang W, Fung TT, Hu FB, Smith-Warner SA, Chavarro JE, et al. Development and validation of empirical indices to assess the insulinaemic potential of diet and lifestyle. Br J Nutr 2016;1–12.
18. Li M, Song L, Yuan J, Zhang D,Zhang C,Liu Y, et al. Association between serum insulin and C-peptide levels and breast cancer: an updated systematic review and meta-analysis. Front Oncol 2020;10:2390.
19. Ahern TP, Hankinson SE, Willett WC, Pollak MN, Eliassen AH, Tamimi RM. Plasma C-peptide, mammographic breast density, and risk of invasive breast cancer. Cancer Epidemiol Biomarkers Prev 2013;22:1786–96.
20. Loopstra-Masters RC, Liese AD, Haffner SM, Wagenknecht LE, Hanley AJ. Associations between the intake of caffeinated and decaffeinated coffee and measures of insulin sensitivity and beta cell function. Diabetologia 2011;54:320–8.
21. Fung TT, Hu FB, Schulze M, Pollak M, Wu T, Fuchs CS, et al. A dietary pattern that is associated with C-peptide and risk of colorectal cancer in women. Cancer Causes Control 2012;23:959–65.
22. Bobe G, Murphy G, Rogers CJ, Hance KW, Albert PS, Laiyemo AO, et al. Serum adiponectin, leptin, C-peptide, homocysteine, and colorectal adenoma recurrence in the polyp prevention trial. Cancer Epidemiol Biomarkers Prev 2010;19:1441–52.
23. Tabung FK, Satija A, Fung TT, Clinton SK, Giovannucci EL. Long-term change in both dietary insulinemic and inflammatory potential is associated with weight gain in adult women and men. J Nutr 2019;149:804–15.
24. Yang W, Sui J, Zhao L, Ma Y, Tabung FK, Simon TG, et al. Association of inflammatory and insulinemic potential of diet and lifestyle with risk of hepatocellular carcinoma. Cancer Epidemiol Biomarkers Prev 2021;30:789–96.
25. Lee DH, Fung TT, Tabung FK, Colditz GA, Ghobrial IM, Rosner BA, et al. Dietary pattern and risk of multiple myeloma in two large prospective US cohort studies. JNCI Cancer Spectr 2019;3:pkz025.
26. Wang W, Fung TT, Wang M, Smith-Warner SA, Giovannucci EL, Tabung FK. Association of the insulinemic potential of diet and lifestyle with risk of digestive system cancers in men and women. JNCI Cancer Spectr 2018;2:pky080.
27. Jin Qi, Shi Ni, Aroke D, Lee DH, Joseph JJ, Donneyong M, et al. Insulinemic and inflammatory dietary patterns show enhanced predictive potential for type 2 diabetes risk in postmenopausal women. Diabetes Care 2021;44:707–14.
28. Lee DH, Li J, Li Y, Liu G, Wu K, Bhupathiraju S, et al. Dietary inflammatory and insulinemic potential and risk of type 2 diabetes: results from three prospective U.S. cohort studies. Diabetes Care 2020;43:2675–83.
29. Tabung FK, Wang W, Fung TT, Smith-Warner SA, Keum N, Wu K, et al. Association of dietary insulinemic potential and colorectal cancer risk in men and women. Am J Clin Nutr 2018;108:363–70.
30. Tabung FK, Noonan A, Lee DH, Song M, Clinton SK, Spakowicz D, et al. Post-diagnosis dietary insulinemic potential and survival outcomes among colorectal cancer patients. BMC Cancer 2020;20:817.
31. Aroke D, Folefac E, Shi Ni, Jin Qi, Clinton SK, Tabung FK. Inflammatory and insulinemic dietary patterns: influence on circulating biomarkers and prostate cancer risk. Cancer Prev Res 2020;13:841–52.
32. Tamimi RM, Baer HJ, Marotti J, Galan M, Galaburda L, Fu Y, et al. Comparison of molecular phenotypes of ductal carcinoma in situ and invasive breast cancer. Breast Cancer Res 2008;10:R67.
33. Michels K, Song M, Willett WC, Rosner B. Latency estimation for chronic disease risk: a damped exponential weighting model. Eur J Epidemiol 2020;35:807–19.
34. Tabung FK, Smith-Warner SA, Chavarro JE, Wu K, Fuchs CS, Hu FB, et al. Development and validation of an empirical dietary inflammatory index. J Nutr 2016;146:1560–70.
35. Lin DY, Fleming TR, De Gruttola V Estimating the proportion of treatment effect explained by a surrogate marker. Stat Med 1997;16:1515–27.
36. Stryker WS, Kaplan LA, Stein EA, Stampfer MJ, Sober A, Willett WC. the relation of diet, cigarette smoking, and alcohol consumption to plasma beta-carotene and alpha-tocopherol levels. Am J Epidemiol 1988;127:283–96.
37. Jun H-J, Austin SB, Wylie SA, Corliss HL, Jackson B, Spiegelman D, et al. The mediating effect of childhood abuse in sexual orientation disparities in tobacco and alcohol use during adolescence: results from the nurses' health study iI. Cancer Causes Control 2010;21: 1817–28.
38. Lunn M, Mcneil D. Applying Cox regression to competing risks. Biometrics 1995;51:524–32.
39. Giovannucci E a framework to understand diet, physical activity, body weight, and cancer risk. Cancer Causes Control 2018;29:1–6.
40. Simone V, Dâ Avenia M, Argentiero A, Felici C, Rizzo FM, De Pergola G, et al. Obesity and breast cancer: molecular interconnections and potential clinical applications. Oncologist 2016;21:404–17.
41. Kang JH, Peng C, Rhee JJ, Farvid MS, Willett WC, Hu FB, et al. Prospective study of a diabetes risk reduction diet and the risk of breast cancer. Am J Clin Nutr. 2020;112:1492–503.
42. Murphy LO, Blenis J. MAPK signal specificity: the right place at the right time. Trends Biochem Sci 2006;31:268–75.
43. Simpson L, Parsons R. Life as a tumor suppressor. Exp Cell Res 2001;264:29–41.
44. Rochefort H, Glondu M, Sahla ME, Platet N, Garcia M. How to target estrogen receptor-negative breast cancer? Endocrine Relat Cancer 2003;10:261–6.
45. Barnard ME, Boeke CE, Tamimi RM. Established breast cancer risk factors and risk of intrinsic tumor subtypes. Biochim Biophys Acta Rev Cancer 2015;1856:73–85.
46. Romieu I, Ferrari P, Rinaldi S, Slimani N, Jenab M, Olsen A, et al. Dietary glycemic index and glycemic load and breast cancer risk in the European Prospective Investigation into Cancer And Nutrition (EPIC). Am J Clin Nutr 2012;96:345–55.
47. Park Y, Brinton LA, Subar AF, Hollenbeck A, Schatzkin A. Dietary fiber intake and risk of breast cancer in postmenopausal women: the National Institutes of Health-AARP Diet and Health Study. Am J Clin Nutr 2009;90:664–71.
48. Farvid MS, Chen WY, Rosner BA, Tamimi RM, Willett WC, Eliassen AH. Fruit and vegetable consumption and breast cancer incidence: repeated measures over 30 years of follow-up. Int J Cancer 2019;144:1496–510.
49. Eliassen AH, Hendrickson SJ, Brinton LA, Buring JE, Campos H, Dai Qi, et al. Circulating carotenoids and risk of breast cancer: pooled

analysis of eight prospective studies. J Natl Cancer Inst 2012;104:1905–16.
50. Sanchez-Quesada C, Romanos-Nanclares A, Navarro AM, Gea A, Cervantes SN, Martinez-Gonzalez MA, et al. Coffee consumption and breast cancer risk in the SUN project. Eur J Nutr 2020;59:3461–71.
51. Romanos-Nanclares A, Willett WC, Rosner BA, Collins LC, Hu FB, Toledo E, et al. Healthful and unhealthful plant-based diets and risk of breast cancer in U.S. women: results from the Nurses' Health Studies. Cancer Epidemiol Biomarkers Prev 2021;30:1921–31.
52. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project expert report 2018. Diet, Nutrition, Physical Activity and Breast Cancer. Available from: dietandcancerreport.org.
53. Fiolet T, Srour B, Sellem L, Kesse-Guyot E, Allés B, MéJean C, et al. Consumption of ultra-processed foods and cancer risk: results from Nutrinet-Santé prospective cohort. BMJ 2018;360:k322.
54. Farvid MS, Stern MC, Norat T, Sasazuki S, Vineis P, Weijenberg MP, et al. Consumption of red and processed meat and breast cancer incidence: a systematic review and meta-analysis of prospective studies. Int J Cancer 2018;143:2787–99.
55. Romanos-Nanclares A, Collins LC, Hu FB, Willett WC, Rosner BA, Toledo E, et al. Sugar-sweetened beverages, artificially sweetened beverages, and breast cancer risk: results from 2 prospective US cohorts. J Nutr 2021;151(9):2768–79.
56. Cohen DH, Leroith D. Obesity, type 2 diabetes, and cancer: the insulin and IGF connection. Endocr Relat Cancer 2012;19:F27–45.
57. Murphy N, Knuppel A, Papadimitriou N, Martin RM, Tsilidis KK, Smith-Byrne K, et al. Insulin-like growth factor-1, insulin-like growth factor-binding protein-3, and breast cancer risk: observational and Mendelian randomization analyses with ∼430 000 women. Ann Oncol 2020;31:641–9.
58. Pollak M. Insulin and insulin-like growth factor signalling in neoplasia. Nat Rev Cancer 2008;8:915–28.
59. Saltiel AR, Kahn CR. Insulin signalling and the regulation of glucose and lipid metabolism. Nature 2001;414:799–806.
60. Engelman JA, Luo Ji, Cantley LC. the evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. Nat Rev Genet 2006;7:606–19.
61. Morita M, Gravel S-P, Chénard V, Sikström K, Zheng L, Alain T, et al. MTORC1 controls mitochondrial activity and biogenesis through 4E-BP-dependent translational regulation. Cell Metab 2013;18:698–711.
62. Senapati P, Thai C, Sanchez A,Leung A,Thai C,Sanchez A, et al. Hyperinsulinemia promotes aberrant histone acetylation in triple negative breast cancer. Epigenetics Chromatin 2019;12:44.
63. Tabung FK, Balasubramanian R, Liang L, Clinton SK, Cespedes Feliciano EM, Manson JE, et al. Identifying metabolomic profiles of insulinemic dietary patterns. Metabolites 2019;9:120.
64. Moore SC. Metabolomics and breast cancer: scaling up for robust results. BMC Med 2020;18:18.
65. His M, Viallon V, Dossus L, Gicquiau A, Achaintre D, Scalbert A, et al. Prospective analysis of circulating metabolites and breast cancer in EPIC. BMC Med 2019;17:178.
66. Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. Int J Epidemiol 1989;18:858–67.
67. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122:51–65.
68. Shi N, Aroke D, Jin Q, Lee DH, Hussan H, Zhang X, et al. Proinflammatory and hyperinsulinemic dietary patterns are associated with specific profiles of biomarkers predictive of chronic inflammation, glucose-insulin dysregulation, and dyslipidemia in postmenopausal women. Front Nutr 2021;20(8):690428.