Prospective study of a diabetes risk reduction diet and the risk of breast cancer

Jae H Kang,[1] Cheng Peng,[1] Jinnie J Rhee,[2] Maryam S Farvid,[3] Walter C Willett,[3,4] Frank B Hu,[1,3] Bernard A Rosner,[1,5] Rulla Tamimi,[1,3] and A Heather Eliassen[1,3]

[1]Channing Division of Network Medicine, Department of Medicine, Brigham & Women's Hospital, Boston, MA, USA; [2]Division of Nephrology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA; [3]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, USA; [4]Department of Nutrition, Harvard TH Chan School of Public Health, Boston, MA, USA; and [5]Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA, USA

ABSTRACT

Background: Hyperinsulinemia and higher insulin-like growth factors may increase breast cancer risk. We evaluated a diabetes risk reduction diet (DRRD) and breast cancer risk.

Objectives: We prospectively evaluated the association between adherence to a DRRD and the incidence of breast cancer.

Methods: We followed 88,739 women from the Nurses' Health Study (NHS; 1980–2016) and 93,915 women from the NHSII (1991–2017). Incident breast cancer cases ($n = 11,943$) were confirmed with medical records, and subtypes were determined by tissue microarray data and pathology reports. Information on diet and breast cancer risk factors was repeatedly ascertained in follow-up questionnaires. A DRRD score was derived with 9 factors: lower glycemic index of diet; lower intakes of *trans* fat, sugar-sweetened beverages/fruit juices, and red/processed meat; higher intakes of cereal fiber, coffee, nuts, and whole fruits; and a higher ratio of polyunsaturated to saturated fat (score range: 9–45). Multivariable-adjusted hazard ratios (MVHRs) and 95% CIs were calculated with Cox proportional hazards models.

Results: Being in the highest compared with the lowest DRRD adherence quintile was associated with a modestly lower breast cancer risk ($\text{MVHR}_{\text{Q5vsQ1}}$: 0.89; 95% CI: 0.84, 0.95; P-trend = 0.0002); this was attenuated after adjusting for weight change since age 18 y ($\text{MVHR}_{\text{Q5vsQ1}}$: 0.92; 95% CI: 0.87, 0.98; P-trend = 0.01). The inverse association was strongest among women with current BMI < 25 kg/m^2 ($\text{MVHR}_{\text{Q5vsQ1}}$: 0.89; 95% CI: 0.81, 0.98; P-trend = 0.004; P-interaction = 0.04). Among tumor molecular subtypes, the strongest inverse association was observed with basal-type tumors ($\text{MVHR}_{\text{Q5vsQ1}}$: 0.67; 95% CI: 0.45, 1.01; P-trend = 0.04).

Conclusions: Greater DRRD-adherence was associated with lower breast cancer risk, likely mediated by less weight gain with a DRRD; however, independently of weight change, DRRD-adherence was modestly associated with lower breast cancer risk, particularly among lean women. *Am J Clin Nutr* 2020;112:1492–1503.

Keywords: breast cancer, epidemiology, diet, risk factors, tumor subtypes

Introduction

Insulin resistance, linked to type 2 diabetes (T2D), has been associated with breast cancer (1–3). Higher C-peptide concentrations, an insulin secretion marker, are associated with elevated breast cancer risk, particularly for estrogen receptor negative (ER-negative) tumors (4). With hyperinsulinemia (a hallmark of insulin resistance), insulin may stimulate cellular signaling pathways involved in growth factor–dependent cell proliferation and cancer development [e.g., microtubule associated protein (MAP) kinase (5) and PI3K/Akt/mTOR (6) pathways]. Insulin increases insulin-like growth factor-1 activity (7), important in
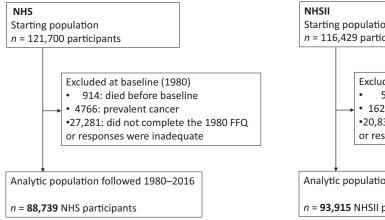
FIGURE 1 Flowchart of study population. NHS, Nurses' Health Study.

tumor initiation and progression (8), and insulin increases estrogen bioavailability (9), which promotes breast carcinogenesis. Thus, encouraging lifestyle modifications to reduce the risk of developing insulin resistance and hyperinsulinemia may be a potential breast cancer primary prevention strategy.

Besides weight maintenance, diet is important in preventing insulin resistance and hyperinsulinemia, and multiple factors may be important. For example, foods high in glycemic index (GI) (10–15) such as sugar-sweetened beverages (SSBs) (16, 17) and refined grains (18–20), red meat (21), and saturated and *trans* fats (22, 23) may increase T2D risk, whereas cereal fiber (24–29), coffee (30–33), nuts (34), polyunsaturated fats (22, 23), and fruits (35, 36) may lower risk. Although some individual factors have been associated with breast cancer (37–39), an overall dietary pattern (40) that emphasizes healthy intake of multiple factors may be more etiologically relevant.

Previously, we developed a score for a diabetes risk reduction diet (DRRD) that emphasizes intake of these multiple factors; it was associated with a 40% lower T2D risk (independent of BMI) (41). Here, we evaluate the hypothesis that greater DRRD adherence is associated with lower breast cancer incidence.

Subjects

The Nurses' Health Study (NHS) is an ongoing study of 121,700 female nurses aged 30–55 y in 1976, and the NHSII has followed 116,429 female nurses (aged 25–42 y at recruitment) since 1989. Every 2 y, participants have provided information on health-related factors and medical history.

Women were followed from 1980 in the NHS and from 1991 in the NHSII, when dietary information was first available, to 2016 in the NHS and to 2017 in the NHSII (**Figure 1**). We excluded women who died before 1980 in the NHS and before 1991 in the NHSII (NHS: 914; NHSII: 54); who had prevalent cancer (NHS: 4766; NHSII: 1621); or who did not return the first diet questionnaires, left >70 food items blank on the first semiquantitative FFQ, or reported implausible total energy intake (<500 or >3500 kcal/d) (NHS: 27,281; NHSII: 20,839), leaving 182,654 women (NHS: 88,739; NHSII: 93,915). The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard TH Chan School of

Public Health, and those of participating state cancer registries as required. The study procedures were in accordance with the ethical standards of the responsible institutional committees.

Methods

Dietary assessment and DRRD score derivation

Diet was assessed with FFQs administered in the NHS in 1980, 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 61 items in 1980, 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 portion size, and participants were asked to specify the food-specific average consumption during the previous year (from among 9 choices ranging from "almost never" to ">6/day"). Participants' nutrient intakes were calculated by multiplying the nutrient content of a food serving (based on updated USDA databases and other sources) and consumption frequency (42, 43). An FFQ validation study reported reasonable correlation between the FFQ and multiple dietary records for coffee (0.78) (44), total and specific types of fat (0.46–0.68) (45, 46), carbohydrates (0.64) (45, 46), fiber (0.56) (45, 46), nuts and peanut butter (0.75) (44), SSBs (0.36–0.84) (44), total fruits (0.70) (44, 47), and red and processed meats (0.38–0.70) (44).

To calculate the DRRD score (DRRDS) (41), we assigned participants a quintile value between 1 (intake consistent with the highest T2D risk) and 5 (for the lowest T2D risk) for each of 9 dietary factors: cereal fiber, nuts, coffee (caffeinated and decaffeinated), whole fruits (raisins, prunes, bananas, cantaloupes, watermelons, fresh apples/pears, oranges, grapefruits, strawberries, blueberries, peaches/apricots/plums), and ratio of polyunsaturated to saturated fat in ascending order; and GI, *trans* fat, SSBs/fruit juices (apple, orange, grapefruit, and other fruit juices), and red and processed meats in descending order. We modified the previous version of the DRRDS (41) by incorporating data on fruits and fruit juices in relation to diabetes risk (35, 36): we added total fruits as a diabetes protective factor and combined fruit juices with SSBs as 1 adverse factor. The DRRDS (range = 9–45) was the sum of the quintile values.

As the FFQ item number evolved over time, we derived FFQ-year-specific quintiles and DRRDSs and calculated cumulatively averaged DRRDSs where the average of all available information before a risk period was used (such averages better represent long-term exposure and have less random measurement error) (48).

Case 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 compared with in situ, and ER, progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) receptor status. For deceased cases, the next of kin was contacted for this 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.

Tissue microarrays, immunohistochemical analysis, and subtype classification (49)

We have previously described details of breast cancer tissue block collection and tissue microarray (TMA) construction (50). 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 5/6 (CK 5/6), and epidermal growth factor receptor (EGFR) (51). In the NHS only (for cases diagnosed between 1980 and 2006), information was available for the insulin receptor (IR): an H score (52) was calculated as percentage of positively stained cells at weak intensity category $\times$ 1 + percentage of positively stained cells at median intensity category $\times$ 2 + percentage of positively stained cells at high intensity category $\times$ 3; IR positivity was defined as H score greater than the median (51). Additional staining for the proliferative marker Ki-67 was completed in NHS cases; Ki-67 data were not available for NHSII cases. Cases with TMAs were very similar to all eligible invasive cases in terms of demographics, breast cancer risk factors, and tumor characteristics.

For tumor molecular subtyping for a subset of cases, we used definitions that correlated with gene expression profile classifications (53–58). For tumors missing Ki-67 expression data (NHSII tumors), histologic grade was used. Thus, 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. Unclassified tumors were ER-negative, PR-negative, HER2-negative, CK 5/6-negative, and EGFR-negative. For evaluating ER-positive compared with ER-negative tumors, ER status was determined primarily from TMA slides, and if unavailable, secondarily from pathology reports.

Statistical analysis

Because we observed no between-cohort heterogeneity (*P*-heterogeneity = 0.30), we pooled the NHS and NHSII data. We allowed eligible participants to contribute person-time from 1980 in the NHS and from 1991 in the NHSII to the breast cancer diagnosis date, date of diagnosis of other cancers (excluding nonmelanoma skin cancers), death, or the end of follow-up [2016 (NHS) or 2017 (NHSII) for the main analysis and 2006 for the molecular subtype analysis], whichever came first. The primary outcome of the analysis was incident breast cancer (occurring in 1980–2016 in the NHS and in 1991–2017 in the NHSII) and secondary outcomes were the various breast cancer tumor subtypes. We assessed the association between DRRDS quintiles and incident breast cancer using multivariable-adjusted time-varying Cox proportional hazards regression models, stratified by age (mo), 2-y time-period at risk, and cohort (NHS or NHSII), to estimate HRs and 95% CIs. We tested for linear trends by evaluating the quintile median values as a continuous variable. Also, we estimated HRs for a 3-point increment in the DRRDS (which was equivalent to the difference in medians across the quintiles of the DRRDS) and its components.

Covariates included race (self-reported), census-tract socioeconomic status variables (median household income, percentage with college degrees, percentage with no high school degree), age at menarche, age at menopause, postmenopausal hormone use, oral contraceptive use, parity, age at first birth, breastfeeding history, height, alcohol intake, total caloric intake, physical activity, and BMI at age 18 y. Because change in weight from age 18 y may be intermediate between diet and breast cancer, we further adjusted for it in a separate model. For variables with missing data, a missing indicator was used and for those with <5% missing data, imputation to the most common category or median value was used.

To evaluate DRRDS individual components (each as cumulatively averaged variables), we evaluated models where all individual factors were entered simultaneously. To test whether the DRRDS and breast cancer association differed by current BMI (in kg/m^2), menopausal status, or diabetes status, we added interaction terms [median DRRDS value across quintiles (continuous variable) $\times$ effect modifier] and used the Wald test. To evaluate the extent to which associations may be potentially mediated by diabetes or weight gain from age 18 y, we performed mediation analyses, where we estimated the mediation proportion (the proportion of the observed association attributable to a mediator) (59).

To evaluate whether associations differed by molecular subtype, we used the Lunn–McNeil approach (60) to derive the *P* value for heterogeneity. *P* values were 2-sided, we used $\alpha = 0.05$, and we performed analyses using SAS version 9.4 (SAS Institute Inc.). Adjustments were not made for multiple comparisons; therefore, the analyses of secondary outcomes and subtypes can be considered exploratory.

DRRDS and breast tumor gene expression

RNA was extracted from multiple cores of 1 or 1.5 mm taken from formalin-fixed paraffin-embedded tumor (*n* = 1–3 cores) and normal-adjacent (*n* = 3–5 cores) tissues using the Qiagen AllPrep RNA isolation kit (50). In brief (61–63), we

profiled transcriptome-wide gene expression using Affymetrix Glue Grant Human Transcriptome Array 3.0 (hGlue 3.0) and Human Transcriptome Array 2.0 (HTA 2.0) microarray chips (Affymetrix). We used the robust multiarray average to perform normalization (Affymetrix Power Tools), log-2 transformed the data, and conducted sample quality control with Affymetrix Power Tools probeset summarization-based metrics (61, 62). From 954 cases, 1577 samples (882 tumor tissues and 695 normal-adjacent tissues) passed quality control. For genes that were mapped by multiple probes, we selected the most variable probe to represent the gene. Genes with low expression (<25th percentile) were removed. We included 17,791 (70%) genes profiled in both platforms. Batch variabilities were controlled using an empirical Bayes method *(Combat)* (64).

Using a competitive gene set testing procedure (CAMERA), we explored functional enrichment of biological pathways associated with DRRDS (65). We chose the 50 "hallmark" gene sets from the Molecular Signatures Database (MSigDB; http://www.broadinstitute.org/gsea/msigdb/) (66) and used the false discovery rate for multiple comparisons (67). We included data from 768 cases with data on DRRDS and gene expression. We chose an intergene correlation of 0.01. All analyses were conducted in SAS version 9.4 (SAS Institute Inc.) and in R version 3.1.4 (R Core Team).

Results

During 4,832,621 person-years of follow-up, we identified 11,943 incident breast cancer cases (NHS, 8027; NHSII, 3916). Women with higher DRRDS were less likely to have diabetes and be African-American; had lower current BMI and gained less weight from age 18 y; and were older and more likely to drink more alcohol, be nulliparous, have breastfed, and be on postmenopausal hormones (**Table 1**).

Although age-adjusted models showed no associations, in multivariable-adjusted models (**Table 2**), the DRRDS was significantly inversely associated with breast cancer, with the model 1 HR, comparing the highest quintile (Q5) with the lowest quintile (Q1), being 0.89 (95% CI: 0.84, 0.95; *P*-trend = 0.0002). When the DRRDS was evaluated as a continuous variable, a 3-point increment in the DRRDS was associated with a 3% lower risk (HR: 0.97; 95% CI: 0.96, 0.98; *P*-trend < 0.0001). Additional adjustment for change in weight since age 18 y slightly attenuated the association (model 2: HR_{Q5vsQ1}: 0.92; 95% CI: 0.87, 0.98; *P*-trend = 0.01; 3-point increment in DRRDS HR: 0.98; 95% CI: 0.97, 0.99; *P*-trend = 0.003) (**Supplemental Figure 1**).

We observed a suggestive interaction by menopausal status (*P*-interaction = 0.09) (Table 2); however, the association was stronger for postmenopausal than for premenopausal breast cancer (model 2: HR_{Q5vsQ1}: 0.90; 95% CI: 0.84, 0.97 for postmenopausal and HR_{Q5vsQ1}: 0.96; 95% CI: 0.84, 1.10 for premenopausal women) (**Supplemental Figure 2**). Among postmenopausal women, there was a suggestion (*P*-interaction = 0.07) of stronger associations among noncurrent users of postmenopausal hormones (model 2: HR_{Q5vsQ1}: 0.85; 95% CI: 0.77, 0.94) than among current users (HR_{Q5vsQ1}: 1.02; 95% CI: 0.89, 1.17).

We evaluated the extent that the inverse association with higher DRRDS in multivariable-adjusted model 1 may be mediated by less weight gain from age 18 y. The calculated mediation proportion was 27.8% (95% CI: 16.0%, 43.8%; *P* < 0.0001), indicating that less weight gain could statistically explain 27.8% of the inverse association with DRRDS. Less weight gain was not a mediating factor for premenopausal breast cancer, but in postmenopausal women, the mediation proportion was 30.1% (95% CI: 17.9%, 46.0%; *P* < 0.0001).

We observed a significant interaction by current BMI (*P*-interaction = 0.04) (**Table 3**), where a significant inverse association with DRRDS was evident in women with BMI < 25, but not among overweight (BMI 25–29) or obese women (BMI ≥30). Model 1 and model 2 (where current BMI was further adjusted for) showed similar results. In model 2, among lean women, the HR_{Q5vsQ1} was 0.89 (95% CI: 0.81, 0.98; *P*-trend = 0.004), which contrasted with the findings in other groups (among overweight women, HR_{Q5vsQ1}: 0.95; 95% CI: 0.85, 1.07; *P*-trend = 0.30; among obese women, HR_{Q5vsQ1}: 0.96; 95% CI: 0.83, 1.11; *P*-trend = 0.99).

The calculated mediation proportion for diabetes in model 1 was 3.4% (95% CI: 1.7%, 6.6%; *P* < 0.0001), indicating that lower diabetes prevalence could statistically explain 3.4% of the inverse association with DRRDS; this attenuated to 1.7% (95% CI: 0.6%, 4.5%; *P* = 0.01) for model 2. We observed no interaction by diabetes status (*P*-interaction = 0.82 for model 2).

We observed no significant heterogeneity by ER status (*P*-heterogeneity = 0.15) (**Table 4**); although inverse associations were suggestively stronger for ER-negative cancer (HR_{Q5vsQ1}: 0.92; 95% CI: 0.78, 1.08; *P*-trend = 0.03). Although *P*-heterogeneity was 0.11 across the subtypes, inverse associations were observed with the HER2-enriched tumors (model 2 HR_{Q5vsQ1}: 0.77; 95% CI: 0.50, 1.18; *P*-trend = 0.05) and basal-like tumors (HR_{Q5vsQ1}: 0.67; 95% CI: 0.45, 1.01; *P*-trend = 0.04). For IR status, the *P*-heterogeneity was 0.36.

For individual DRRDS components (**Supplemental Table 1**), weak inverse associations were observed with total coffee (HR_{Q5vsQ1}: 0.96; 95% CI: 0.90, 1.01; *P*-trend = 0.11) and whole fruits (HR_{Q5vsQ1}: 0.92; 95% CI: 0.86, 0.99; *P*-trend = 0.07). For ER-negative tumors, total coffee intake (HR_{Q5vsQ1}: 0.85; 95% CI: 0.73, 0.99; *P*-trend = 0.05) was inversely associated. For basal-like tumors, the strongest associations were observed with total coffee (HR_{Q5vsQ1}: 0.65; 95% CI: 0.45, 0.94; *P*-trend = 0.02) and *trans* fat intake (HR_{Q5vsQ1}: 1.65; 95% CI: 1.05, 2.58; *P*-trend = 0.05).

In breast tumor tissue gene expression analyses (*n* = 768), with multivariable-adjusted model 1, 5 biological pathways were significantly downregulated with higher DRRDS (**Table 5**), which included 2 immune-regulatory pathways (interferon *γ* response and interferon *α* response) and 3 pathways related to proliferation (allograft rejection, mTOR signaling, and E2F targets). With adjustment for change in weight since age 18 y, with higher DRRDS, the E2F response was no longer significant (likely due to higher weight being associated with higher E2F pathway activation) (68, 69) whereas 4 pathways remained significantly downregulated, and a new proliferation-related pathway (Myc targets v2) was also downregulated.

Discussion

In this study of 180,000 women followed for ≥26 y, women with greatest adherence to a 9-item DRRDS showed a modestly lower risk of incident invasive breast cancer, which

TABLE 1 Age and age-standardized characteristics of participants by quintiles of DRRDS as of the midpoint of follow-up (1996 in the NHS and 1997 in the NHSII)[1]

	NHS		NHSII	
	Q1	Q5	Q1	Q5
n	15,431	15,300	19,342	18,925
Age, y	60.3 ± 7.0	64.0 ± 6.9	41.8 ± 4.7	43.6 ± 4.4
DRRDS	21.1 ± 1.9	33.1 ± 1.9	19.4 ± 2.3	34.2 ± 2.2
Ratio of polyunsaturated fat to saturated fat[2]	0.4 ± 0.1	0.6 ± 0.3	0.4 ± 0.1	0.6 ± 0.2
Cereal fiber,[2] g/d	3.2 ± 1.3	5.2 ± 2.2	4.2 ± 1.5	7.7 ± 3.4
Total coffee intake,[2] cups/d	1.5 ± 1.4	2.8 ± 1.5	0.7 ± 1.2	2.2 ± 1.6
Total nut or peanut butter intake,[2] servings/d	0.2 ± 0.2	0.5 ± 0.5	0.1 ± 0.2	0.3 ± 0.4
Total whole fruit intake, servings/d	0.9 ± 0.6	2.2 ± 1.0	0.6 ± 0.5	1.7 ± 0.9
Glycemic index of diet[2]	54.5 ± 2.5	50.3 ± 2.6	55.7 ± 2.5	52.0 ± 2.6
Trans fat intake,[2] % of total kcal/d	2.0 ± 0.5	1.4 ± 0.4	1.9 ± 0.5	1.1 ± 0.4
Sugar-sweetened beverage/fruit juice intake,[2] servings/d	1.4 ± 1.0	0.7 ± 0.6	1.7 ± 1.4	0.7 ± 0.7
Total red meat intake,[2] servings/d	1.5 ± 0.6	0.7 ± 0.4	1.3 ± 0.6	0.5 ± 0.4
Total caloric intake,[2] kcal/d	1743.5 ± 434.2	1669.6 ± 408.5	1838.1 ± 517.8	1790.6 ± 477.5
Total vegetable intake,[2] servings/d	3.1 ± 1.3	4.8 ± 1.9	2.5 ± 1.3	4.7 ± 2.4
Alcohol intake,[2] g/d	5.3 ± 9.4	5.9 ± 8.2	2.1 ± 4.9	4.1 ± 6.2
Self-reported history of diabetes	8.4	5.5	2.3	1.7
Current BMI, kg/m[2]	27.1 ± 5.7	26.0 ± 4.9	27.1 ± 6.8	25.3 ± 5.3
BMI at age 18 y, ≥22 kg/m[2]	28.0	36.6	27.3	32.7
Weight change from age 18 y, kg	16.4 ± 13.8	11.5 ± 12.2	16.6 ± 14.6	10.4 ± 12.4
Height, in	64.4 ± 2.4	64.6 ± 2.4	64.8 ± 2.6	65.0 ± 2.6
Physical activity, metabolic-equivalent-of-task-h/wk	12.9 ± 18.1	23.4 ± 25.3	13.4 ± 18.2	25.6 ± 28.2
Self-reported African heritage	2.1	1.4	2.6	0.9
Census-tract median annual family income, $	60,484.8 ± 21,081.2	68,669.0 ± 26,165.6	57,877.9 ± 18,358.0	67,431.8 ± 23,425.3
Family history of breast cancer	13.0	13.9	8.6	9.1
Personal history of benign breast disease	44.5	49.9	15.5	15.4
Age at menarche < 12 y	20.8	25.0	22.6	26.5
Ever used oral contraceptives	50.5	51.2	87.4	86.2
Nulliparous	5.1	6.7	17.1	26.3
Parity,[3] n	3.2 ± 1.6	3.0 ± 1.4	2.3 ± 0.9	2.2 ± 0.9
Ever breastfed[3]	58.0	70.7	74.3	86.9
Postmenopausal	87.7	88.5	12.1	10.3
Age at menopause,[4] y	47.3 ± 5.4	47.8 ± 5.3	37.9 ± 3.7	38.0 ± 4.2
Current postmenopausal hormone use[5]	42.2	53.6	75.2	73.0

[1] Values are means ± SD.s or percentages and are standardized to the age distribution of the study population; given the long follow-up (≥26 y), the characteristics at follow-up midpoint were selected to better represent the study population characteristics. DRRDS, diabetes risk reduction diet score; NHS, Nurses' Health Study; Q, quintile. The SI equivalent of 1 cup is 250 ml, and 1 inch is quivalent to 2.54 cm.

[2] Intakes were adjusted for total energy and represent cumulatively updated intakes.

[3] Among parous women only.

[4] Among women with natural menopause or bilateral oophorectomy.

[5] Among postmenopausal women.

was independent of weight change, and this association was most evident in lean women and in relation to basal-like and HER2-enriched tumors. Because this was the first evaluation of a DRRD and breast cancer, confirmation is needed.

Adiposity plays a key role in breast cancer, with higher concentration of adipocytokines and hyperinsulinemia/insulin resistance promoting breast cancer cell growth (70) and with greater estradiol production in postmenopausal women (71, 72). Indeed, lowering the likelihood of weight gain from age 18 y explained 27% of the inverse association with DRRD-adherence, particularly among postmenopausal women. However, a mechanism independent of adiposity [e.g., stimulating cell proliferation pathways such as MAP kinase (5) and PI3K/Akt/mTOR (6) pathways] is likely given the inverse association evident after adjustment for weight change and the identified pathways in gene expression analyses. Indeed, whereas among overweight/obese

women, we observed no DRRDS association, among lean women, greater DRRDS was inversely associated with risk, even after adjusting for weight change. This is consistent with observations that unhealthy diets were most strongly associated with risk of T2D in lean women (73).

Our finding of stronger inverse DRRDS associations with ER-negative breast cancers (for which nonhormonal exposures may be most important) (74–76) is consistent with other studies that have reported lower risks of ER-negative breast cancer associated with lower glycemic load (77), lower total carbohydrate intake (77), higher dietary fiber (78), higher fruits and vegetable intake (79–81), and greater adherence to either a Dietary Approaches to Stop Hypertension (82, 83) or a Mediterranean diet (84). Among the DRRDS components, we observed the strongest associations with coffee. In a meta-analysis, a weak inverse association with breast cancer was observed (85), whereas in

TABLE 2 Multivariable HRs and 95% CIs for the association between quintiles of cumulatively updated DRRDS and breast cancer using pooled data from the NHS (follow-up from 1980–2016) and NHSII (follow-up from 1991–2017)[1]

	Quintiles of DRRDS					P-trend	Association with each 3-point increase in score
	1	2	3	4	5		
Median score	21	24	27	30	33		
All breast cancer (case n = 11,943)							
Cases/person-years	2336/973,798	2385/970,420	2375/958,612	2416/949,397	2431/980,394		
Age-adjusted model	1.00 (ref)	1.00 (0.94, 1.06)	0.98 (0.92, 1.03)	0.98 (0.92, 1.04)	0.95 (0.90, 1.00)	0.05	0.99 (0.97, 1.00); P = 0.02
Multivariable model 1[2]	1.00 (ref)	0.97 (0.92, 1.03)	0.95 (0.89, 1.00)	0.93 (0.87, 0.99)	0.89 (0.84, 0.95)	0.0002	0.97 (0.96, 0.98); P < 0.0001
Multivariable model 2[3]	1.00 (ref)	0.98 (0.92, 1.03)	0.96 (0.90, 1.01)	0.95 (0.89, 1.00)	0.92 (0.87, 0.98)	0.01	0.98 (0.97, 0.99); P = 0.003
Premenopausal breast cancer (case n = 2561)[4]							
Cases	536	530	508	510	477		
Age-adjusted model	1.00 (ref)	1.02 (0.90, 1.15)	1.01 (0.89, 1.14)	1.05 (0.93, 1.18)	0.97 (0.85, 1.10)	0.72	0.99 (0.97, 1.01); P = 0.32
Multivariable model 1[2]	1.00 (ref)	1.01 (0.90, 1.15)	1.01 (0.89, 1.14)	1.04 (0.92, 1.18)	0.96 (0.84, 1.11)	0.68	0.98 (0.96, 1.01); P = 0.26
Multivariable model 2[3]	1.00 (ref)	1.01 (0.89, 1.14)	1.00 (0.88, 1.14)	1.04 (0.91, 1.18)	0.96 (0.84, 1.10)	0.67	0.98 (0.96, 1.01); P = 0.25
Postmenopausal breast cancer (case n = 8619)[4]							
Cases	1622	1708	1721	1767	1801		
Age-adjusted model	1.00 (ref)	1.00 (0.93, 1.07)	0.97 (0.90, 1.04)	0.95 (0.89, 1.02)	0.93 (0.87, 0.99)	0.01	0.98 (0.97, 1.00); P = 0.01
Multivariable model 1[2]	1.00 (ref)	0.97 (0.90, 1.04)	0.93 (0.87, 1.00)	0.90 (0.84, 0.96)	0.86 (0.80, 0.93)	<0.0001	0.96 (0.95, 0.98); P < 0.0001
Multivariable model 2[3]	1.00 (ref)	0.98 (0.91, 1.04)	0.95 (0.88, 1.01)	0.92 (0.86, 0.99)	0.90 (0.84, 0.97)	0.003	0.97 (0.96, 0.99); P = 0.002
P-interaction by menopausal status[5] = 0.09							

[1]DRRDS, diabetes risk reduction diet score; NHS, Nurses' Health Study.

[2]Multivariable model 1 from Cox regression analyses with time-varying covariates: stratified by cohort, age (mo), and 2-y period at risk, adjusted for race (non-Hispanic Caucasian, African, Asian, Hispanic Caucasian), quintiles of the percentage of those aged ≥25 y in the residential census-tract that are college educated; quintiles of the percentage of those aged ≥25 y in the residential census-tract without a high school degree, quintiles of the median family income of the residential census-tract; age at menarche (<12, 12, 13, 14, >14 y), age at menopause (premenopausal, <45, 45–49, 50–52, ≥53 y), 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 y, current user of other types), oral contraceptive use history (never, past, current use), parity and age at first birth (nulliparous, 1 child before age 25 y, 1 child at ≥25 y of age, ≥2 children before age 25 y, ≥2 children at ≥25 y of age), breastfeeding history (never, breastfed for ≤6 mo, breastfed for >6 mo), family history of breast cancer (yes, no), history of 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 (quintiles; kcal/d), cumulatively updated total vegetable intake (linear; servings/d), physical activity (linear; metabolic-equivalent-of-task-h/wk), and BMI at age 18 y (<20.0, 20.0–21.9, 22.0–23.9, 24.0–26.9, ≥27.0).

[3]Multivariable model 2 = Multivariable model 1 + change in weight since age 18 y (lost ≥2, lost 0–1, gained 0–2, gained 3–5, gained 6–10, gained 11–20, gained 21–25, gained >25 kg).

[4]The person-time of women whose menopausal status was unclear was excluded.

[5]Wald test of an interaction term of menopausal status × continuous variable representing quintile median values of DRRDS in multivariable model 2.

TABLE 3 Multivariable HRs and 95% CIs for the association between quintiles of cumulatively updated DRRDS and breast cancer using pooled data from the NHS (follow-up from 1980–2016) and NHSII (follow-up from 1991–2017): by current BMI[1]

| | Quintiles of DRRDS | | | | | P-trend | Association with each 3-point increase in score |
	1	2	3	4	5		
Median score	21	24	27	30	33		
BMI < 25 (case n = 5269)							
Cases, n	942	1011	1014	1073	1229		
Age-adjusted model	1.00 (ref)	0.98 (0.90, 1.07)	0.96 (0.88, 1.05)	0.94 (0.86, 1.02)	0.94 (0.86, 1.02)	0.08	0.98 (0.97, 1.00); P = 0.06
Multivariable model 1[2]	1.00 (ref)	0.96 (0.87, 1.05)	0.93 (0.84, 1.01)	0.89 (0.81, 0.97)	0.88 (0.80, 0.97)	0.003	0.97 (0.95, 0.99); P = 0.002
Multivariable model 2[3]	1.00 (ref)	0.95 (0.87, 1.04)	0.93 (0.84, 1.01)	0.89 (0.81, 0.97)	0.89 (0.81, 0.98)	0.004	0.97 (0.95, 0.99); P = 0.003
BMI = 25–29 (case n = 3688)							
Cases, n	686	771	736	778	717		
Age-adjusted model	1.00 (ref)	1.04 (0.94, 1.15)	0.97 (0.88, 1.08)	1.01 (0.91, 1.12)	0.93 (0.83, 1.03)	0.12	0.98 (0.96, 1.00); P = 0.09
Multivariable model 1[2]	1.00 (ref)	1.04 (0.94, 1.16)	0.98 (0.88, 1.09)	1.01 (0.90, 1.12)	0.95 (0.84, 1.07)	0.28	0.99 (0.96, 1.01); P = 0.24
Multivariable model 2[3]	1.00 (ref)	1.04 (0.94, 1.16)	0.98 (0.88, 1.09)	1.01 (0.90, 1.13)	0.95 (0.85, 1.07)	0.30	0.99 (0.96, 1.01); P = 0.27
BMI ≥ 30 (case n = 2580)							
Cases, n	607	538	540	498	397		
Age-adjusted model	1.00 (ref)	0.94 (0.84, 1.06)	1.00 (0.89, 1.13)	1.01 (0.89, 1.13)	0.97 (0.85, 1.10)	0.91	0.99 (0.97, 1.02); P = 0.68
Multivariable model 1[2]	1.00 (ref)	0.93 (0.82, 1.04)	0.99 (0.88, 1.12)	0.99 (0.87, 1.12)	0.95 (0.83, 1.10)	0.93	0.99 (0.96, 1.02); P = 0.48
Multivariable model 2[3]	1.00 (ref)	0.93 (0.82, 1.05)	0.99 (0.88, 1.12)	0.99 (0.87, 1.13)	0.96 (0.83, 1.11)	0.99	0.99 (0.96, 1.02); P = 0.54
P-interaction by BMI[4] = 0.04							

[1]The person-time of women with missing BMI was excluded. BMI in kg/m^2. DRRDS, diabetes risk reduction diet score; NHS, Nurses' Health Study.
[2]As per footnote 2 of Table 2, except for oral contraceptive use (never, ever).
[3]Multivariable model 2 = multivariable model 1 + current BMI (linear).
[4]Wald test of an interaction term of continuous variable representing median values of the 3 BMI categories × continuous variable representing quintile median values of DRRDS in Multivariable model 1.

TABLE 4 Multivariable HRs and 95% CIs for the association between quintiles of cumulatively updated DRRDS and breast cancer tumor subtypes in the NHS and NHSII[1]

	DRRDS quintiles, median					P-trend	Association with each 3-point increase in score
	Q1, 21	Q2, 24	Q3, 27	Q4, 30	Q5, 33		
By ER status (NHS with follow-up of 1980–2016 and NHSII with follow-up of 1991–2017)							
ER positive breast cancer (case n = 7678)							
Cases	1445	1528	1512	1590	1603		
Multivariable model 1[2]	1.00 (ref)	0.99 (0.92, 1.07)	0.95 (0.88, 1.02)	0.96 (0.89, 1.04)	0.93 (0.86, 1.00)	0.05	0.98 (0.96, 0.99); P = 0.01
Multivariable model 2[3]	1.00 (ref)	1.00 (0.93, 1.07)	0.96 (0.89, 1.04)	0.98 (0.91, 1.06)	0.96 (0.88, 1.04)	0.29	0.99 (0.97, 1.00); P = 0.10
ER negative breast cancer (case n = 1778)							
Cases	387	397	343	304	347		
Multivariable model 1[2]	1.00 (ref)	1.02 (0.88, 1.17)	0.90 (0.78, 1.05)	0.79 (0.67, 0.92)	0.90 (0.77, 1.06)	0.01	0.96 (0.93, 0.99); P = 0.01
Multivariable model 2[3]	1.00 (ref)	1.02 (0.88, 1.18)	0.91 (0.78, 1.06)	0.79 (0.68, 0.93)	0.92 (0.78, 1.08)	0.03	0.96 (0.93, 1.00); P = 0.03
P-heterogeneity by ER status = 0.15[4]							
By molecular subtype (NHS with follow-up of 1980–2006 and NHSII with follow-up of 1991–2005)							
Luminal A (case n = 2810)							
Cases	518	546	544	599	603		
Multivariable model 1[5]	1.00 (ref)	0.99 (0.87, 1.12)	0.96 (0.85, 1.08)	1.02 (0.90, 1.16)	0.97 (0.85, 1.11)	0.85	0.99 (0.97, 1.02); P = 0.60
Multivariable model 2[6]	1.00 (ref)	0.99 (0.88, 1.12)	0.97 (0.85, 1.09)	1.03 (0.91, 1.17)	0.99 (0.87, 1.13)	0.90	1.00 (0.97, 1.02); P = 0.85
Luminal B (case n = 1204)							
Cases	224	271	244	229	236		
Multivariable model 1[5]	1.00 (ref)	1.17 (0.98, 1.40)	1.04 (0.86, 1.25)	0.95 (0.78, 1.16)	0.98 (0.80, 1.20)	0.38	0.97 (0.93, 1.01); P = 0.12
Multivariable model 2[6]	1.00 (ref)	1.18 (0.99, 1.41)	1.06 (0.88, 1.27)	0.98 (0.80, 1.19)	1.02 (0.83, 1.25)	0.68	0.98 (0.94, 1.02); P = 0.29
HER2-enriched (case n = 265)							
Cases	57	69	43	48	48		
Multivariable model 1[5]	1.00 (ref)	1.15 (0.80, 1.64)	0.70 (0.47, 1.06)	0.76 (0.51, 1.14)	0.77 (0.50, 1.18)	0.05	0.91 (0.83, 0.99); P = 0.04
Multivariable model 2[6]	1.00 (ref)	1.14 (0.80, 1.63)	0.70 (0.47, 1.05)	0.76 (0.51, 1.14)	0.77 (0.50, 1.18)	0.05	0.91 (0.83, 0.99); P = 0.04
Basal-like (case n = 315)							
Cases	72	74	64	60	45		
Multivariable model 1[5]	1.00 (ref)	1.04 (0.75, 1.45)	0.94 (0.66, 1.33)	0.86 (0.60, 1.23)	0.64 (0.42, 0.96)	0.02	0.92 (0.85, 0.99); P = 0.03
Multivariable model 2[6]	1.00 (ref)	1.06 (0.76, 1.47)	0.96 (0.68, 1.36)	0.89 (0.62, 1.27)	0.67 (0.45, 1.01)	0.04	0.93 (0.86, 1.00); P = 0.05
Unclassified (case n = 92)							
Cases	16	18	23	12	23		
Multivariable model 1[5]	1.00 (ref)	1.14 (0.57, 2.24)	1.56 (0.81, 2.99)	0.77 (0.36, 1.68)	1.44 (0.72, 2.87)	0.57	1.06 (0.91, 1.22); P = 0.45
Multivariable model 2[6]	1.00 (ref)	1.14 (0.58, 2.25)	1.57 (0.82, 3.03)	0.78 (0.36, 1.69)	1.47 (0.73, 2.95)	0.53	1.06 (0.92, 1.23); P = 0.41
P-heterogeneity by molecular subtype = 0.11[4,7]							
By IR status (NHS only with follow-up of 1980–2006)							
IR positive breast cancer[8] (case n = 1123)							
Cases	227	220	220	246	210		
Multivariable model 1[5]	1.00 (ref)	0.94 (0.78, 1.13)	0.93 (0.77, 1.12)	0.99 (0.82, 1.19)	0.83 (0.67, 1.02)	0.13	0.96 (0.92, 1.01); P = 0.09
Multivariable model 2[6]	1.00 (ref)	0.94 (0.78, 1.13)	0.93 (0.77, 1.13)	0.99 (0.82, 1.20)	0.84 (0.68, 1.03)	0.17	0.97 (0.92, 1.01); P = 0.12
IR negative breast cancer[8] (case n = 1159)							
Cases	225	249	224	226	235		
Multivariable model 1[5]	1.00 (ref)	1.08 (0.90, 1.30)	0.98 (0.81, 1.19)	0.93 (0.77, 1.13)	0.97 (0.79, 1.18)	0.38	0.99 (0.94, 1.03); P = 0.50
Multivariable model 2[6]	1.00 (ref)	1.09 (0.91, 1.31)	0.99 (0.82, 1.20)	0.94 (0.78, 1.14)	0.99 (0.81, 1.21)	0.52	0.99 (0.95, 1.03); P = 0.67
P-heterogeneity by IR status = 0.36[4]							

[1]DRRDS, diabetes risk reduction diet score; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; IR, insulin receptor; NHS, Nurses' Health Study; Q, quintile.

[2]As per footnote 2 of Table 2, except oral contraceptive use (never, ever).

[3]As per footnote 3 of Table 2, except oral contraceptive use (never, ever).

[4]For testing heterogeneity by subtype, we used the Lunn–McNeil approach (60), for multivariable model 2.

[5]Owing to smaller sample sizes in analyses, to ensure that models would run, covariate categorizations were simplified. Multivariable model 1: stratified by cohort, age (mo), and 2-y period at risk, adjusted for age at menarche (<12, 12, 13, ≥14 y), age at menopause (premenopausal, <45, 45–49, ≥50 y), postmenopausal hormone use (nonuser, current user—estrogen only, current user + progestin user, current user of other types), oral contraceptive use history (never, ever), parity and age at first birth (nulliparous, 1 child only, ≥2 children before age 25 y, ≥2 children ≥ 25 y of age), breastfeeding history (never, breastfed for ≤6 mo, breastfed for >6 mo), family history of breast cancer (yes, no), history of benign breast disease (yes, no), height (<1.60, 1.60–1.64, 1.65–1.69, ≥1.70 m), cumulatively updated alcohol intake (linear; g/d), cumulatively updated total caloric intake (linear; kcal/d), cumulatively updated total vegetable intake (linear; servings/d), physical activity (linear; metabolic-equivalent-of-task-h/wk), and BMI (in kg/m²) at age 18 y (<20.0, 20.0–21.9, 22.0–23.9, 24.0–26.9, ≥27.0).

[6]Multivariable model 2 = multivariable model 1 + change in weight since age 18 y (linear; kg).

[7]In post hoc analyses to test for heterogeneity by subtype, the basal-like type was significantly different from luminal A type (P for difference = 0.047; for comparison with luminal B type, P = 0.11; HER2-enriched type, P = 0.91; unclassified type, P = 0.13).

[8]IR positive or negative status was determined as ≥ (positive) or < (negative) the median of IR expression (cytoplasmic and membranous).

TABLE 5 Multivariable-adjusted gene expression pathway analysis showing associations with higher cumulatively updated DRRDS (modeled as a continuous variable) in breast tumor tissue among 768 cases[1]

	Genes, *n*	Direction of gene expression regulation	*P* value	FDR
Pathways identified with multivariable-adjusted model 1[2]				
HALLMARK_INTERFERON_GAMMA_RESPONSE (immune regulation)	157	Down	1.42E-06	3.58E-05
HALLMARK_INTERFERON_ALPHA_RESPONSE (immune regulation)	78	Down	1.43E-06	3.58E-05
HALLMARK_ALLOGRAFT_REJECTION (proliferation)	152	Down	3.20E-05	5.33E-04
HALLMARK_MTORC1_SIGNALING (proliferation)	170	Down	8.38E-04	1.05E-02
HALLMARK_E2F_TARGETS (proliferation)	151	Down	1.90E-03	1.90E-02
Pathways identified with multivariable-adjusted model 2[3]				
HALLMARK_INTERFERON_ALPHA_RESPONSE (immune regulation)	78	Down	1.82E-05	9.12E-04
HALLMARK_INTERFERON_GAMMA_RESPONSE (immune regulation)	157	Down	4.46E-05	1.11E-03
HALLMARK_ALLOGRAFT_REJECTION (proliferation)	152	Down	8.88E-04	1.48E-02
HALLMARK_MTORC1_SIGNALING (proliferation)	170	Down	3.05E-03	3.51E-02
HALLMARK_MYC_TARGETS_V2 (proliferation)	49	Down	3.51E-03	3.51E-02

[1]DRRDS, diabetes risk reduction diet score; FDR, false discovery rate.

[2]We used a competitive gene set testing procedure to explore the functional enrichment of biological pathways (65) associated with the DRRDS. Multivariable-adjusted model 1 adjusted for age at diagnosis, year of diagnosis, estrogen receptor status, menopausal status, physical activity, alcohol consumption, total calorie intake, family history of breast cancer, history of benign breast disease, BMI at age 18 y, and total vegetable intake.

[3]Multivariable-adjusted model 2 further adjusted for change in weight since age 18 y.

some studies, like ours, the strongest inverse associations were observed with ER-negative breast cancers (86, 87), potentially due to inhibition of tumorigenesis by antioxidant polyphenols (88) and influences on hormone concentrations (89). Among the ER-negative breast cancers, higher DRRDS was associated with 23%–33% lower risks for HER2-enriched and basal-like tumors. Factors related to adiposity/hyperinsulinemia may be important (75) for these subtypes. Triple negative breast cancers (TNBCs), of which 80% are basal-like tumors, are known to be especially sensitive to the anticancer effects of metformin (an insulin-sensitizing drug). This effect of metformin may operate through inhibition of the mTOR pathway (90); mTOR pathway deregulation is relatively prominent in TNBCs (91) and is associated with poor outcomes among TNBC patients (92, 93). For HER2-enriched cancers, some studies (94, 95) have shown metformin to be inversely associated with risk. Metformin can suppress HER2 protein overexpression (96) and delay cancer onset in animal models (97), and it was associated with reduced risk of fatal HER2-positive breast cancer (95). Thus, although there is some mechanistic support for the associations with breast cancer subtypes, further confirmatory studies are warranted.

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. Limitations include inevitable measurement errors in assessing diets, which would likely be nondifferential in relation to risk of breast cancer; these may have caused underestimation of associations. However, the DRRDS has previously been strongly associated with lower T2D risk (41), indicating that the score is well measured. One limitation was that we could not evaluate diet from childhood/adolescence, which may be a critical period for breast cancer initiation (98–100). Also, we had limited power to evaluate molecular subtypes. In addition, residual confounding cannot be excluded, but we controlled for a wide variety of risk factors, and the similar socioeconomic background of participants helped to minimize this potential bias. Finally, because the participants were all

highly trained medical professionals who were mostly Caucasian, the results may not be generalizable to populations with different underlying breast cancer risks.

In sum, higher DRRD adherence was associated with lower overall breast cancer risk, with some of this association being explained by the lower weight gain observed with DRRD adherence. However, even independently of weight change, higher adherence to a DRRD was modestly associated with lower breast cancer risk, particularly among lean women and postmenopausal women. Further studies of this dietary pattern and breast cancer are warranted.

References

1. Pasanisi P, Berrino F, De Petris M, Venturelli E, Mastroianni A, Panico S. Metabolic syndrome as a prognostic factor for breast cancer recurrences. Int J Cancer 2006;119(1):236–8.
2. Bruning PF, Bonfrer JM, van Noord PA, Hart AA, de Jong-Bakker M, Nooijen WJ. Insulin resistance and breast-cancer risk. Int J Cancer 1992;52(4):511–16.
3. Giovannucci E, Harlan DM, Archer MC, Bergenstal RM, Gapstur SM, Habel LA, Pollak M, Regensteiner JG, Yee D. Diabetes and cancer: a consensus report. Diabetes Care 2010;33(7):1674–85.
4. 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(10):1786–96.
5. Murphy LO, Blenis J. MAPK signal specificity: the right place at the right time. Trends Biochem Sci 2006;31(5):268–75.

6. Simpson L, Parsons R. PTEN: life as a tumor suppressor. Exp Cell Res 2001;264(1):29–41.

7. Straus DS. Nutritional regulation of hormones and growth factors that control mammalian growth. FASEB J 1994;8(1):6–12.

8. Pollak M. Insulin and insulin-like growth factor signalling in neoplasia. Nat Rev Cancer 2008;8(12):915–28.

9. Stephenson GD, Rose DP. Breast cancer and obesity: an update. Nutr Cancer 2003;45(1):1–16.

10. Augustin LS, Franceschi S, Jenkins DJ, Kendall CW, La Vecchia C. Glycemic index in chronic disease: a review. Eur J Clin Nutr 2002;56(11):1049–71.

11. Barclay AW, Petocz P, McMillan-Price J, Flood VM, Prvan T, Mitchell P, Brand-Miller JC. Glycemic index, glycemic load, and chronic disease risk—a meta-analysis of observational studies. Am J Clin Nutr 2008;87(3):627–37.

12. Frost G, Leeds A, Trew G, Margara R, Dornhorst A. Insulin sensitivity in women at risk of coronary heart disease and the effect of a low glycemic diet. Metabolism 1998;47(10):1245–51.

13. Liu S, Manson JE, Buring JE, Stampfer MJ, Willett WC, Ridker PM. Relation between a diet with a high glycemic load and plasma concentrations of high-sensitivity C-reactive protein in middle-aged women. Am J Clin Nutr 2002;75(3):492–8.

14. Ludwig DS. The glycemic index: physiological mechanisms relating to obesity, diabetes, and cardiovascular disease. JAMA 2002;287(18):2414–23.

15. Wu T, Giovannucci E, Pischon T, Hankinson SE, Ma J, Rifai N, Rimm EB. Fructose, glycemic load, and quantity and quality of carbohydrate in relation to plasma C-peptide concentrations in US women. Am J Clin Nutr 2004;80(4):1043–9.

16. Hu FB, Malik VS. Sugar-sweetened beverages and risk of obesity and type 2 diabetes: epidemiologic evidence. Physiol Behav 2010;100(1):47–54.

17. Schulze MB, Manson JE, Ludwig DS, Colditz GA, Stampfer MJ, Willett WC, Hu FB. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. JAMA 2004;292(8):927–34.

18. Aune D, Norat T, Romundstad P, Vatten LJ. Whole grain and refined grain consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of cohort studies. Eur J Epidemiol 2013;28(11):845–58.

19. Hu EA, Pan A, Malik V, Sun Q. White rice consumption and risk of type 2 diabetes: meta-analysis and systematic review. BMJ 2012;344:e1454.

20. Sun Q, Spiegelman D, van Dam RM, Holmes MD, Malik VS, Willett WC, Hu FB. White rice, brown rice, and risk of type 2 diabetes in US men and women. Arch Intern Med 2010;170(11):961–9.

21. The InterAct Consortium, Bendinelli B, Palli D, Masala G, Sharp SJ, Schulze MB, Guevara M, van der A D, Sera F, Amiano P, et al. Association between dietary meat consumption and incident type 2 diabetes: the EPIC-InterAct study. Diabetologia 2013;56(1):47–59.

22. de Souza RJ, Mente A, Maroleanu A, Cozma AI, Ha V, Kishibe T, Uleryk E, Budylowski P, Schünemann H, Beyene J, et al. Intake of saturated and trans unsaturated fatty acids and risk of all cause mortality, cardiovascular disease, and type 2 diabetes: systematic review and meta-analysis of observational studies. BMJ 2015;351:h3978.

23. Salmeron J, Hu FB, Manson JE, Stampfer MJ, Colditz GA, Rimm EB, Willett WC. Dietary fat intake and risk of type 2 diabetes in women. Am J Clin Nutr 2001;73(6):1019–26.

24. Andersson A, Tengblad S, Karlström B, Kamal-Eldin A, Landberg R, Basu S, Åman P, Vessby B. Whole-grain foods do not affect insulin sensitivity or markers of lipid peroxidation and inflammation in healthy, moderately overweight subjects. J Nutr 2007;137(6):1401–7.

25. Karupaiah T, Aik CK, Heen TC, Subramaniam S, Bhuiyan AR, Fasahat P, Zain AM, Ratnam W. A transgressive brown rice mediates favourable glycaemic and insulin responses. J Sci Food Agric 2011;91(11):1951–6.

26. Panlasigui LN, Thompson LU. Blood glucose lowering effects of brown rice in normal and diabetic subjects. Int J Food Sci Nutr 2006;57(3–4):151–8.

27. Pereira MA, Jacobs DR Jr, Pins JJ, Raatz SK, Gross MD, Slavin JL, Seaquist ER. Effect of whole grains on insulin sensitivity in overweight hyperinsulinemic adults. Am J Clin Nutr 2002;75(5):848–55.

28. Cho SS, Qi L, Fahey GC Jr, Klurfeld DM. Consumption of cereal fiber, mixtures of whole grains and bran, and whole grains and risk reduction in type 2 diabetes, obesity, and cardiovascular disease. Am J Clin Nutr 2013;98(2):594–619.

29. Okarter N, Liu RH. Health benefits of whole grain phytochemicals. Crit Rev Food Sci Nutr 2010;50(3):193–208.

30. Bhupathiraju SN, Pan A, Manson JE, Willett WC, van Dam RM, Hu FB. Changes in coffee intake and subsequent risk of type 2 diabetes: three large cohorts of US men and women. Diabetologia 2014;57(7):1346–54.

31. Ding M, Bhupathiraju SN, Chen M, van Dam RM, Hu FB. Caffeinated and decaffeinated coffee consumption and risk of type 2 diabetes: a systematic review and a dose-response meta-analysis. Diabetes Care 2014;37(2):569–86.

32. Doo T, Morimoto Y, Steinbrecher A, Kolonel LN, Maskarinec G. Coffee intake and risk of type 2 diabetes: the Multiethnic Cohort. Public Health Nutr 2014;17(6):1328–36.

33. Huxley R, Lee CM, Barzi F, Timmermeister L, Czernichow S, Perkovic V, Grobbee DE, Batty D, Woodward M. Coffee, decaffeinated coffee, and tea consumption in relation to incident type 2 diabetes mellitus: a systematic review with meta-analysis. Arch Intern Med 2009;169(22):2053–63.

34. Zhou D, Yu H, He F, Reilly KH, Zhang J, Li S, Zhang T, Wang B, Ding Y, Xi B. Nut consumption in relation to cardiovascular disease risk and type 2 diabetes: a systematic review and meta-analysis of prospective studies. Am J Clin Nutr 2014;100(1):270–7.

35. Muraki I, Imamura F, Manson JE, Hu FB, Willett WC, van Dam RM, Sun Q. Fruit consumption and risk of type 2 diabetes: results from three prospective longitudinal cohort studies. BMJ 2013;347:f5001.

36. Bazzano LA, Li TY, Joshipura KJ, Hu FB. Intake of fruit, vegetables, and fruit juices and risk of diabetes in women. Diabetes Care 2008;31(7):1311–17.

37. Mullie P, Koechlin A, Boniol M, Autier P, Boyle P. Relation between breast cancer and high glycemic index or glycemic load: a meta-analysis of prospective cohort studies. Crit Rev Food Sci Nutr 2016;56(1):152–9.

38. Guo J, Wei W, Zhan L. Red and processed meat intake and risk of breast cancer: a meta-analysis of prospective studies. Breast Cancer Res Treat 2015;151(1):191–8.

39. Yu X, Bao Z, Zou J, Dong J. Coffee consumption and risk of cancers: a meta-analysis of cohort studies. BMC Cancer 2011;11:96.

40. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002;13(1):3–9.

41. Rhee JJ, Mattei J, Hughes MD, Hu FB, Willett WC. Dietary diabetes risk reduction score, race and ethnicity, and risk of type 2 diabetes in women. Diabetes Care 2015;38(4):596–603.

42. Willett WC, Sampson L, Browne ML, Stampfer MJ, Rosner BA, Hennekens CH, Speizer FE. The use of a self-administered questionnaire to assess diet four years in the past. Am J Epidemiol 1988;127(1):188–99.

43. Yuan C, Spiegelman D, Rimm EB, Rosner BA, Stampfer MJ, Barnett JB, Chavarro JE, Subar AF, Sampson LK, Willett WC. Validity of a dietary questionnaire assessed by comparison with multiple weighed dietary records or 24-hour recalls. Am J Epidemiol 2017;185(7):570–84.

44. Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, Willett WC. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. Int J Epidemiol 1989;18(4):858–67.

45. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer FE. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122(1):51–65.

46. Willett WC. Nutritional epidemiology. 2nd ed. New York: Oxford University Press; 1998.

47. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB, Willett WC. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. J Am Diet Assoc 1993;93(7):790–6.

48. Hu FB, Stampfer MJ, Rimm E, Ascherio A, Rosner BA, Spiegelman D, Willett WC. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. Am J Epidemiol 1999;149(6):531–40.

49. Colditz GA, Rosner BA, Chen WY, Holmes MD, Hankinson SE. Risk factors for breast cancer according to estrogen and progesterone receptor status. J Natl Cancer Inst 2004;96(3):218–28.

50. 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.

51. Wang J, Zhang X, Beck AH, Collins LC, Chen WY, Tamimi RM, Hazra A, Brown M, Rosner B, Hankinson SE. Alcohol consumption and risk of breast cancer by tumor receptor expression. Horm Cancer 2015;6(5–6):237–46.

52. Kinsel LB, Szabo E, Greene GL, Konrath J, Leight GS, McCarty KS Jr. Immunocytochemical analysis of estrogen receptors as a predictor of prognosis in breast cancer patients: comparison with quantitative biochemical methods. Cancer Res 1989;49(4):1052–6.

53. Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, Hernandez-Boussard T, Livasy C, Cowan D, Dressler L, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 2004;10(16):5367–74.

54. Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, Perou CM, Nielsen TO. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. Clin Cancer Res 2008;14(5):1368–76.

55. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 2009;101(10):736–50.

56. van de Rijn M, Perou CM, Tibshirani R, Haas P, Kallioniemi O, Kononen J, Torhorst J, Sauter G, Zuber M, Köchli OR, et al. Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. Am J Pathol 2002;161(6):1991–6.

57. Brenton JD, Carey LA, Ahmed AA, Caldas C. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? J Clin Oncol 2005;23(29):7350–60.

58. Abd El-Rehim DM, Pinder SE, Paish CE, Bell J, Blamey RW, Robertson JF, Nicholson RI, Ellis IO. Expression of luminal and basal cytokeratins in human breast carcinoma. J Pathol 2004;203(2):661–71.

59. Lin DY, Fleming TR, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. Stat Med 1997;16(13):1515–27.

60. Lunn M, McNeil D. Applying Cox regression to competing risks. Biometrics 1995;51(2):524–32.

61. Wang J, Heng YJ, Eliassen AH, Tamimi RM, Hazra A, Carey VJ, Ambrosone CB, de Andrade VP, Brufsky A, Couch FJ, et al. Alcohol consumption and breast tumor gene expression. Breast Cancer Res 2017;19(1):108.

62. Heng YJ, Wang J, Ahearn TU, Brown SB, Zhang X, Ambrosone CB, de Andrade VP, Brufsky AM, Couch FJ, King TA, et al. Molecular mechanisms linking high body mass index to breast cancer etiology in post-menopausal breast tumor and tumor-adjacent tissues. Breast Cancer Res Treat 2018;173:667–77.

63. Kensler KH, Sankar VN, Wang J, Zhang X, Rubadue CA, Baker GM, Parker JS, Hoadley KA, Stancu AL, Pyle ME, et al. PAM50 molecular intrinsic subtypes in the Nurses' Health Study cohorts. Cancer Epidemiol Biomarkers Prev 2019;28(4):798–806.

64. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007;8(1):118–27.

65. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res 2012;40(17):e133.

66. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1(6):417–25.

67. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statist Soc B 1995;57(1):289–300.

68. Fajas L, Landsberg RL, Huss-Garcia Y, Sardet C, Lees JA, Auwerx J. E2Fs regulate adipocyte differentiation. Dev Cell 2002;3(1):39–49.

69. Docquier A, Augereau P, Lapierre M, Harmand PO, Badia E, Annicotte JS, Fajas L, Cavailles V. The *RIP140* gene is a transcriptional target of E2F1. PLoS One 2012;7(5):e35839.

70. Simone V, D'Avenia M, Argentiero A, Felici C, Rizzo FM, De Pergola G, Silvestris F. Obesity and breast cancer: molecular interconnections and potential clinical applications. Oncologist 2016;21(4):404–17.

71. Siiteri PK. Adipose tissue as a source of hormones. Am J Clin Nutr 1987;45(1 Suppl):277–82.

72. Grodin JM, Siiteri PK, MacDonald PC. Source of estrogen production in postmenopausal women. J Clin Endocrinol Metab 1973;36(2):207–14.

73. Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, Willett WC. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. N Engl J Med 2001;345(11):790–7.

74. Sisti JS, Collins LC, Beck AH, Tamimi RM, Rosner BA, Eliassen AH. Reproductive risk factors in relation to molecular subtypes of breast cancer: results from the nurses' health studies. Int J Cancer 2016;138(10):2346–56.

75. Barnard ME, Boeke CE, Tamimi RM. Established breast cancer risk factors and risk of intrinsic tumor subtypes. Biochim Biophys Acta 2015;1856(1):73–85.

76. Martinez ME, Cruz GI, Brewster AM, Bondy ML, Thompson PA. What can we learn about disease etiology from case-case analyses? Lessons from breast cancer. Cancer Epidemiol Biomarkers Prev 2010;19(11):2710–14.

77. Romieu I, Ferrari P, Rinaldi S, Slimani N, Jenab M, Olsen A, Tjonneland A, Overvad K, Boutron-Ruault MC, Lajous M, 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(2):345–55.

78. 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(3):664–71.

79. Jung S, Spiegelman D, Baglietto L, Bernstein L, Boggs DA, van den Brandt PA, Buring JE, Cerhan JR, Gaudet MM, Giles GG, et al. Fruit and vegetable intake and risk of breast cancer by hormone receptor status. J Natl Cancer Inst 2013;105(3):219–36.

80. Fung TT, Chiuve SE, Willett WC, Hankinson SE, Hu FB, Holmes MD. Intake of specific fruits and vegetables in relation to risk of estrogen receptor-negative breast cancer among postmenopausal women. Breast Cancer Res Treat 2013;138(3):925–30.

81. 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(7):1496–510.

82. 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.

83. Hirko KA, Willett WC, Hankinson SE, Rosner BA, Beck AH, Tamimi RM, Eliassen AH. Healthy dietary patterns and risk of breast cancer by molecular subtype. Breast Cancer Res Treat 2016;155(3):579–88.

84. 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(2):466–72.

85. Jiang W, Wu Y, Jiang X. Coffee and caffeine intake and breast cancer risk: an updated dose-response meta-analysis of 37 published studies. Gynecol Oncol 2013;129(3):620–9.

86. Lowcock EC, Cotterchio M, Anderson LN, Boucher BA, El-Sohemy A. High coffee intake, but not caffeine, is associated with reduced estrogen receptor negative and postmenopausal breast cancer risk with no effect modification by CYP1A2 genotype. Nutr Cancer 2013;65(3):398–409.

87. Bhoo-Pathy N, Peeters PH, Uiterwaal CS, Bueno-de-Mesquita HB, Bulgiba AM, Bech BH, Overvad K, Tjønneland A, Olsen A, Clavel-Chapelon F, et al. Coffee and tea consumption and risk of pre- and postmenopausal breast cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort study. Breast Cancer Res 2015;17:15.

88. Ramos S. Cancer chemoprevention and chemotherapy: dietary polyphenols and signalling pathways. Mol Nutr Food Res 2008;52(5):507–26.

89. Sisti JS, Hankinson SE, Caporaso NE, Gu F, Tamimi RM, Rosner B, Xu X, Ziegler R, Eliassen AH. Caffeine, coffee, and tea intake and

urinary estrogens and estrogen metabolites in premenopausal women. Cancer Epidemiol Biomarkers Prev 2015;24(8):1174–83.

90. Amin S, Lux A, O'Callaghan F. The journey of metformin from glycaemic control to mTOR inhibition and the suppression of tumour growth. Br J Clin Pharmacol 2019;85(1):37–46.

91. Walsh S, Flanagan L, Quinn C, Evoy D, McDermott EW, Pierce A, Duffy MJ. mTOR in breast cancer: differential expression in triple-negative and non-triple-negative tumors. Breast 2012;21(2):178–82.

92. Ueng S-H, Chen S-C, Chang Y-S, Hsueh S, Lin Y-C, Chien H-P, Lo Y-F, Shen S-C, Hsueh C. Phosphorylated mTOR expression correlates with poor outcome in early-stage triple negative breast carcinomas. Int J Clin Exp Pathol 2012;5(8):806–13.

93. Bayraktar S, Hernandez-Aya LF, Lei X, Meric-Bernstam F, Litton JK, Hsu L, Hortobagyi GN, Gonzalez-Angulo AM. Effect of metformin on survival outcomes in diabetic patients with triple receptor-negative breast cancer. Cancer 2012;118(5):1202–11.

94. Sonnenblick A, Agbor-Tarh D, Bradbury I, Di Cosimo S, Azim HA Jr, Fumagalli D, Sarp S, Wolff AC, Andersson M, Kroep J, et al. Impact of diabetes, insulin, and metformin use on the outcome of patients with human epidermal growth factor receptor 2–positive primary breast cancer: analysis from the ALTTO Phase III randomized trial. J Clin Oncol 2017;35(13):1421–9.

95. He X, Esteva FJ, Ensor J, Hortobagyi GN, Lee MH, Yeung SC. Metformin and thiazolidinediones are associated with improved breast cancer-specific survival of diabetic women with HER2+ breast cancer. Ann Oncol 2012;23(7):1771–80.

96. Vazquez-Martin A, Oliveras-Ferraros C, Menendez JA. The antidiabetic drug metformin suppresses HER2 (erbB-2) oncoprotein overexpression via inhibition of the mTOR effector p70S6K1 in human breast carcinoma cells. Cell Cycle 2009;8(1):88–96.

97. Anisimov VN, Egormin PA, Piskunova TS, Popovich IG, Tyndyk ML, Yurova MN, Zabezhinski MA, Anikin IV, Karkach AS, Romanyukha AA. Metformin extends life span of HER-2/neu transgenic mice and in combination with melatonin inhibits growth of transplantable tumors in vivo. Cell Cycle 2010;9(1):188–97.

98. Linos E, Willett WC, Cho E, Frazier L. Adolescent diet in relation to breast cancer risk among premenopausal women. Cancer Epidemiol Biomarkers Prev 2010;19(3):689–96.

99. Liu Y, Colditz GA, Cotterchio M, Boucher BA, Kreiger N. Adolescent dietary fiber, vegetable fat, vegetable protein, and nut intakes and breast cancer risk. Breast Cancer Res Treat 2014;145(2):461–70.

100. Farvid MS, Cho E, Chen WY, Eliassen AH, Willett WC. Adolescent meat intake and breast cancer risk. Int J Cancer 2015;136(8):1909–20.