Dietary patterns and breast cancer risk in the California Teachers Study cohort[1–5]

Lilli B Link, Alison J Canchola, Leslie Bernstein, Christina A Clarke, Daniel O Stram, Giske Ursin, and Pamela L Horn-Ross

ABSTRACT

Background: Evidence that diet is associated with breast cancer risk is inconsistent. Most studies have examined risks associated with specific foods and nutrients, rather than measures of overall diet.

Objective: This study aimed to evaluate dietary patterns and their relation to breast cancer risk in a large cohort of women.

Design: Data from 91,779 women in the California Teachers Study cohort were analyzed, including data from 4140 women with a diagnosis of invasive breast cancer made between 1995 and 2009. Five predominant dietary patterns were identified by using principal components factor analysis: a plant-based diet, high in fruit and vegetables; a high-protein, high-fat diet, high in meats, eggs, fried foods, and high-fat condiments; a high-carbohydrate diet, high in convenience foods, pasta, and bread products; an ethnic diet, high in legumes, soy-based foods, rice, and dark-green leafy vegetables; and a salad and wine diet, high in lettuce, fish, wine, low-fat salad dressing, and coffee and tea.

Results: The plant-based pattern was associated with a reduction in breast cancer risk (RR: 0.85; 95% CI: 0.76, 0.95 for the highest compared with the lowest consumption quintile; *P*-trend = 0.003); risk reduction was greater for estrogen receptor–negative progesterone receptor–negative (ER–PR–) tumors (RR: 0.66; 95% CI: 0.48, 0.91; *P*-trend = 0.03). The salad and wine pattern was associated with an increased risk of estrogen receptor–positive progesterone receptor–positive tumors (RR: 1.29; 95% CI: 1.12, 1.49); this effect was only slightly attenuated after adjustment for alcohol consumption.

Conclusion: The finding that greater consumption of a plant-based dietary pattern is associated with a reduced breast cancer risk, particularly for ER−PR− tumors, offers a potential avenue for prevention. *Am J Clin Nutr* 2013;98:1524–32.

INTRODUCTION

Epidemiologic evidence of a relation between diet and breast cancer risk is inconsistent; only alcohol consumption has been shown repeatedly to affect risk (1, 2). One explanation is that most analyses have focused on individual foods or nutrients and have not accounted for the complex interactions that occur between the various nutrients and nonnutritive components of different foods (3, 4). These interactions may only be captured statistically when overall measures of diet are analyzed. Thus, evaluating diet as a whole, based on dietary patterns, may provide additional information regarding risk and have important implications for public health recommendations because people do not eat nutrients or foods in isolation.

The use of statistical methods that quantify overall diet has increased over time (5–21). Many previous studies have used principal components factor analysis (PCFA)[6] to identify dietary patterns and to assess the relation between dietary patterns and cancer risk. Most studies have identified at least a healthy and unhealthy dietary pattern, and some studies have identified additional variations of these patterns. For breast cancer, a drinker dietary pattern, high in alcohol consumption, has been associated with increased risk (6, 13, 14, 17), as has a Western or meat-based pattern in some (7, 13, 14, 17, 18, 21, 22) but not all studies (11, 12, 14–16). Diets high in fruit, vegetables, and salad (5, 12, 16, 18); a southern diet, high in cooked greens, legumes, and sweet potatoes (20); a Mediterranean diet (17, 23, 24); and

[1] From Cancer Epidemiology, Mailman School of Public Health, Columbia University, New York, NY (LBL); the Cancer Prevention Institute of California, Fremont, CA (AJC, CAC, and PLH-R); the Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA (CAC and PLH-R); the Division of Cancer Etiology, Department of Population Sciences, Beckman Research Institute, City of Hope, Duarte, CA (LB); the Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA (DOS and GU); the Cancer Registry of Norway, Oslo, Norway (GU); and the Department of Nutrition, University of Oslo, Oslo, Norway (GU).

[2] The ideas and opinions expressed herein are those of the authors, and endorsement by the CDHS, NCI, and CDC or their contractors and subcontractors is not intended nor should be inferred.

[3] Supported in part by grants R01 CA77398 and R25 CA094061 from the National Cancer Institute and by contract 97-10500 from the California Breast Cancer Research Fund. The collection of cancer incidence data used in this study was supported by the California Department of Health Services (CDHS) as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the NCI's Surveillance, Epidemiology and End Results Program under contract N01-PC-35136 awarded to the Northern California Cancer Center, contract N01-PC-35139 awarded to the University of Southern California, and contract N02-PC-15105 awarded to the Public Health Institute (PHI); and the CDC National Program of Cancer Registries under agreement U55/CCR921930-02 awarded to the PHI.

[4] Current address of LBL: Clinical Epidemiology and Evaluative Sciences Research, Weill Cornell Medical College, New York, NY.

[5] Address correspondence and reprint requests to PL Horn-Ross, Cancer Prevention Institute of California, 2201 Walnut Avenue, Suite 300, Fremont, CA. E-mail: pam@cpic.org.

[6] Abbreviations used: CTS, California Teachers Study; ER, estrogen receptor; ER–, estrogen receptor negative; ER+, estrogen receptor positive; HT, hormone therapy; PCFA, principal components factor analysis; PR, progesterone receptor; PR–, progesterone receptor negative; PR+, progesterone receptor positive.

Received February 20, 2013. Accepted for publication September 25, 2013.
First published online October 9, 2013; doi: 10.3945/ajcn.113.061184.

a prudent dietary pattern, high in fruit, vegetables, whole grains, and fish (8, 11, 14, 21) have also been associated with a reduced risk. Other studies, however, have found no significant relations (9, 10, 15, 19). In addition, whether these associations are modified by hormonal factors or vary by hormone-receptor subtype of the cancer is currently not clear.

The purpose of this study was to evaluate the relation between dietary patterns and risk of breast cancer and its subtypes, defined by hormone receptor status, by using PCFA in the large and diverse California Teachers Study (CTS) cohort.

SUBJECTS AND METHODS

The CTS cohort comprises 133,479 active and retired female teachers and administrators who completed a 16-page mailed questionnaire in 1995–1996 (25). The cohort is followed annually for cancer diagnosis, change of address, and death. Cancer diagnoses are determined by linkage with the California Cancer Registry, a population-based cancer registry that covers the entire state of California. Because >99% of all cancer diagnoses among California residents are reported to the California Cancer Registry, those cohort members who continue to reside in California are actively followed for cancer outcomes without the need for further contact. Changes of address are obtained by annual mailings, notifications from participants, and record linkages with multiple sources, including the US Postal Service. California and national mortality files are used to ascertain date and cause of death.

For the current analysis, we excluded women sequentially from the total cohort of 133,479 women if they did not reside in California at baseline ($n = 8867$); they received their diagnosis of breast cancer before joining the cohort ($n = 6211$) or their history of breast cancer was not known ($n = 139$); they had missing or incomplete dietary or alcohol data ($n = 9302$); their self-reported food consumption averaged >3500 or <800 calories/d ($n = 6726$); their age was ≥85 y at baseline ($n = 1294$); they reported >50% of calories from alcohol ($n = 42$); they had unknown vitamin use ($n = 2$); or they had missing data for one or more confounding factors included in this analysis [other than menopausal status/hormone therapy (HT) use] ($n = 9117$). Of the 91,779 women included in this analysis, 4140 received a diagnosis of invasive breast cancer after joining the cohort, with follow-up ending 31 December 2009. Estrogen receptor (ER) and progesterone receptor (PR) status was available for 3483 (84%) of the women with a diagnosis of invasive breast cancer.

The baseline questionnaire gathered information about potential breast cancer risk factors and personal and lifestyle characteristics, including menstrual and reproductive events; use of exogenous estrogens, vitamins, and medications; personal and family history of cancer and chronic diseases; physical activity; self-reported height and weight; dietary intake; and consumption of alcohol and tobacco use. An early version of the 103-item, 1995 Block food-frequency questionnaire, was administered to obtain dietary intake over the year before baseline and has been validated in this cohort (26). A summary measure of neighborhood socioeconomic status was determined by linking the baseline residential street addresses of cohort members to US census data at the block group level; this measure incorporates neighborhood data on occupation, education, and income (27).

The CTS study was approved by the Institutional Review Boards at all participating institutions. This secondary data analysis was approved by the Institutional Review Board of the Cancer Prevention Institute of California (formerly the Northern California Cancer Center) and Columbia University.

Dietary patterns

PCFA was used to identify dietary patterns based on the entire cohort (not just the women included in the current analysis), who were aged ≤84 y at baseline, who completed the dietary assessment, and whose average caloric intake was between 600 and 5000 kcal/d (28). PCFA is a method of reducing numerous variables, in this case foods, beverages (including separate items for wine, beer, and liquor), and vitamin supplements, into fewer variables, or principal components, each of which represents a group of foods and beverages that are correlated with each other but not with the other components. PCFA was chosen for this study because it uses all the dietary information from the food-frequency questionnaire, and, unlike the statistical methods used in reduced rank regression, it does not require the use of an intermediate endpoint. For this analysis, intake of each food or beverage was calculated as the portion size–adjusted frequency of consumption and expressed as the number of medium servings consumed per day. Each variable was standardized to a mean of zero and variance of one, and then a coefficient was assigned to each variable, depending on the weight it contributed to that principal component. Factor scores for each subject for each dietary pattern were calculated by multiplying the participant's consumption of each food or beverage by the weight (ie, factor loading) given to that food or beverage for that dietary pattern and summing these values for all the foods and beverages. An orthogonal varimax rotation procedure was used to obtain uncorrelated components. The number of factors (components) retained was based on eigenvalues ≥1, examination of scree plots, and factor interpretability (29). PCFA using oblique promax rotation, excluding women with a prior breast cancer, any prior cancer, or calorie-adjusting foods and beverages produced the same 5 factors with extremely similar factor loadings.

Data analysis

We used Cox proportional hazards regression to estimate the RRs (hazard rate ratio) and 95% CIs associated with the development of invasive breast cancer for each of the dietary patterns. All analyses used age (in d) as the time metric and were stratified by age at baseline (in y). Each woman was given a factor score for each dietary pattern. Factor scores were categorized into quintiles; higher scores represent greater adherence to that dietary pattern. Factor scores for all 5 dietary patterns for each woman were included in each model.

Women were followed from the date they completed the baseline questionnaire until diagnosis with invasive or in situ breast cancer, death, a move out of California, or 31 December 2009, whichever occurred first. We tested the assumption of proportional hazards for each covariate using a likelihood ratio test of interaction with the time metric (continuous age) based on cross-product terms. Only the effect for a history of benign breast biopsy changed significantly with age (being more strongly associated with risk among younger than among older women); thus, an interaction

term with time-dependent age was included as an adjustment factor. Models were adjusted for race-ethnicity/birthplace (non-Latina white, other race-ethnicity born in North America, other race-ethnicity born elsewhere); family history of breast cancer in a first-degree relative (mother or sister; yes, no, adopted); age at menarche (in y from ≤9 to ≥17); parity (nulliparous, parous), and, among parous women, age at first full-term pregnancy (in y); average daily caloric intake (kcal); average annual long-term (high school to age 54 y or age at baseline if younger) moderate physical activity (h/wk); neighborhood-level socioeconomic status (deciles based on the statewide distribution); history of a benign breast biopsy (yes, no) and its interaction with time-dependent age; BMI at baseline (in kg/m^2); height at baseline (in inches); and menopausal status at baseline [premenopausal, perimenopausal, postmenopausal, undetermined (ie, women aged <56 y who started HT while still having menses or who reported a simple hysterectomy), unknown], and among perimenopausal and postmenopausal women, HT use at baseline (never used, past use only, current use of estrogen-alone therapy, current use of combined estrogen plus progesterone therapy). Likelihood ratio tests for trend across dietary pattern quintiles were conducted by using an ordinal variable coded as the median value of each category. Effect modification by menopausal status/HT use [premenopausal, postmenopausal/not currently using HT (ie, never or past use), postmenopausal/current use of HT] and by BMI (in kg/m^2; <25, ≥25] was examined, as was modification by calendar period (1995–2002, 2003–2009 corresponding to the large cessation of HT use after media coverage of the Women's Health Initiative) for associations among current HT users. Interactions were formally evaluated by using a likelihood ratio test comparing models with and without cross-product terms. However, because none of these interactions were statistically significant, these stratified results are not presented.

For analyses stratified by ER and PR status, all women with a known ERPR subtype (ie, ER+PR+, ER+PR−, ER−PR+, and ER−PR− subtypes) were included in each model, with women with a diagnosis of a different ERPR subtype than the one being censored at the time of their diagnosis; for example, in the analysis of ER+PR+ breast cancer, women with ER+PR−, ER−PR+, and ER−PR− tumors contributed person-time to the analysis until the date of their breast cancer diagnosis. Women with invasive breast cancer whose ER or PR status was borderline or unknown were excluded from these analyses. There were 2422 ER+PR+ cases, 509 ER+PR− cases, and 514 ER−PR− cases included in these analyses. ER−PR+ breast cancers were not analyzed separately because of the small number of cases with this subtype (n = 38). We used SAS (version 9.3; SAS Institute Inc) for all analyses.

RESULTS

We identified 5 major dietary patterns and designated them as plant-based; high-protein, high-fat; high-carbohydrate; ethnic; and salad and wine (**Table 1**). The plant-based pattern was characterized by a high consumption of fruit and vegetables. The high-protein, high-fat dietary pattern included sources of animal protein (eg, meat and eggs) and added fats (eg, butter and mayonnaise). The high-carbohydrate pattern was characterized by a high consumption of convenience foods, pasta, and bread. The ethnic pattern was high in legumes, soy-based foods, rice, and dark-green leafy vegetables. The salad and wine pattern was characterized by a high consumption of salad and low-fat dressing, fish, wine, and coffee and tea. Together, these 5 factors explained 18.6% of the total variation in standardized dietary intakes in the CTS cohort (plant-based: 4.7%; high-protein, high-fat: 4.5%; high-carbohydrate: 3.3%; ethnic: 3.2%, and salad and wine: 2.9%). Fifteen percent of the eligible cohort primarily consumed a single dietary pattern (defined as having a score greater than the median for only one factor), whereas 82% had diets characterized by a combination of patterns; only 3% had diets that were not characterized by any of the 5 patterns.

The characteristics of the women included in this analysis overall and those scoring in the highest quintile of each of the 5 dietary patterns are described in **Table 2**. Median cohort follow-up was 14.1 y (IQR: 13.6, 14.1). Although the median and IQR for age at baseline varied by dietary pattern, the absolute age ranged from 22 to 23–84 y for all dietary patterns. Whereas 36% of the women in the analytic cohort scored in the top quintile of only one pattern, the mixed diet consumed by most of the cohort members resulted in 28% scoring in the top quintile of more than one dietary pattern and 36% not scoring in the top quintile of any of the dietary patterns. Of the 64% of women who scored in the highest quintile of one or more of the dietary patterns, the women in the highest quintile of the plant-based or salad and wine pattern were generally older, more likely to be white, more likely to have reported a history of a benign breast biopsy, and less likely to be nulliparous. Those in the highest quintile of the salad and wine pattern also were more likely to live in neighborhoods with a higher socioeconomic status and were more likely to use HT. Women in the highest quintile of the high-protein, high-fat or high-carbohydrate dietary pattern consumed the greatest amount of calories. Those in the highest quintile of the high-protein, high-fat pattern were also heavier, less likely to engage in moderate physical activity, and of lower socioeconomic status. Women in the highest quintile of the high-carbohydrate pattern were younger and more likely to be nulliparous and of moderate socioeconomic status. Finally, those in the highest quintile of the ethnic dietary pattern were more likely to be non-white, nulliparous, and of lower socioeconomic status.

The plant-based diet was associated with an overall reduction in breast cancer risk, whereas the salad and wine diet was associated with an increased risk (**Table 3**). These overall effects were generally modest, with the highest quintile of intake of the plant-based diet associated with a 15% lower breast cancer risk (95% CI: 0.76, 0.95) relative to the lowest quintile and the salad and wine diet with a 12 greater risk (95% CI: 1.01, 1.25) for the highest relative to the lowest quintile. Nonetheless, a statistically significant trend was evident for both patterns. The increased risk associated with the salad and wine pattern was only slightly related to its alcohol component. The correlation between alcohol consumption and the score for the salad and wine pattern was 0.44. Adjustment for alcohol intake [none, <20 g/d, ≥20 g/d; categories based on our previous findings (30)] reduced the RR associated with the highest compared with the lowest quintile of this pattern to 1.09 (95% CI: 0.97, 1.23; P-trend = 0.06). Furthermore, the RR associated with alcohol consumption was not completely accounted for by consumption of a salad and wine dietary pattern: the RR for alcohol consumption of ≥20 g/d (compared with none) was reduced from 1.24 (95% CI: 1.11, 1.38) to 1.12 (95% CI: 0.99, 1.27) with adjustment for all 5 dietary

TABLE 1
Factor loadings from principal components factor analysis for foods and beverages defining dietary patterns in the California Teachers Study, $n = 118,465$[1]

Food item	Dietary pattern				
	Plant-based	High-protein, high-fat	High-carbohydrate	Ethnic	Salad and wine
Peaches, apricots (fresh)	0.57	—	—	—	—
Strawberries, other berries	0.55	—	—	—	—
Carrots, mixed vegetables with carrots	0.52	—	—	—	—
Apples, apple sauce	0.51	—	—	—	—
Other fruit (eg, grapes, raisins, kiwi, fruit cocktail)	0.50	—	—	—	—
Oranges	0.46	—	—	—	—
Broccoli	0.46	—	—	—	—
Bananas	0.45	—	—	—	—
Watermelon	0.44	—	—	—	—
Cantaloupe	0.41	—	—	—	—
Other vegetables	0.39	—	—	—	—
String beans, green beans	0.38	—	—	—	—
Peas	0.36	—	—	—	—
Cauliflower, Brussels sprouts	0.36	—	—	—	—
Peaches, apricots (dried)	0.35	—	—	—	—
Butter, margarine, or fat added to vegetables	—	0.55	—	—	—
Beef roasts, steaks, sandwiches	—	0.52	—	—	—
Sausage, bacon	—	0.49	—	—	—
Pork	—	0.46	—	—	—
Hamburgers, cheeseburgers	—	0.45	—	—	—
Fried chicken	—	0.44	—	—	—
Beef stew or pot pie with vegetables	—	0.43	—	—	—
Eggs	—	0.40	—	—	—
Fried potatoes	—	0.39	0.36	—	—
Butter on bread or rolls	—	0.39	—	—	—
Salad dressing or mayonnaise (regular)	—	0.38	—	—	—
Burritos or tacos with meat or beans	—	—	0.54	—	—
Pizza	—	—	0.50	—	—
Salsa, ketchup, taco sauce	—	—	0.49	—	—
Tortillas	—	—	0.43	—	—
Spaghetti, lasagna, other pasta with tomato sauce	—	—	0.42	—	—
Bagels, English muffins, hamburger buns	—	—	0.40	—	—
Lentil, pea, and bean soups	—	—	—	0.50	—
Beans (eg, baked beans, pinto, kidney, but not including in soup)	—	—	—	0.49	—
Tofu, bean curd	—	—	—	0.44	—
Vegetable soups	—	—	—	0.39	—
Rice	—	—	—	0.38	—
Meat substitutes made from soy	—	—	—	0.38	—
Mustard, turnip greens, collards	—	—	—	0.37	—
Sweet potatoes, yams	—	—	—	0.35	—
Green salad	—	—	—	—	0.50
Fish (broiled or baked)	—	—	—	—	0.43
Wine, champagne	—	—	—	—	0.42
Salad dressing or mayonnaise (low-fat)	—	—	—	—	0.42
Coffee, tea	—	—	—	—	0.37
Tomatoes, tomato juice	—	—	—	—	0.37

[1] Foods and beverages with factor loadings ≥0.35 are shown.

patterns, including the salad and wine pattern. The plant-based dietary pattern was less affected by alcohol intake. The correlation between alcohol intake and the score for the plant-based pattern was −0.23. Adjustment for alcohol intake resulted in an RR of 0.88 (95% CI: 0.78, 0.99) for consumption of the highest compared with the lowest quintile of the plant-based dietary pattern. The risks associated with the salad and wine (RR: 1.10; 95% CI: 0.96, 1.27 for the highest compared with the lowest quintile) and the plant-based (RR: 0.80; 95% CI: 0.69, 0.93) dietary patterns among those exclusively consuming each of these patterns (ie, excluding women who scored in the highest quintile of more than one dietary pattern) were similar to those observed in all women (as reported in Table 3).

The associations between dietary pattern and breast cancer by hormone receptor status are presented in **Table 4**. The plant-based diet was associated with a reduction in risk of ER−PR− tumors (RR: 0.66; 95% CI: 0.48, 0.91 for the highest compared with the lowest quintile; P-trend = 0.03), whereas the salad and

TABLE 2
Baseline characteristics of all 91,779 participants and of those scoring in the highest quintile of each of the 5 dietary patterns[1]

		Dietary pattern[2]				
Characteristic	All participants	Plant-based	High-protein, high-fat	High-carbohydrate	Ethnic	Salad and wine
Age (y)	50 (42, 60)[3]	55 (46, 66)	53 (44, 64)	44 (36, 50)	51 (42, 61)	57 (49, 66)
Race-ethnicity and birthplace (%)						
Non-Latina white	88	92	87	89	77	94
Nonwhite, born in North America	10	7	11	9	17	5
Nonwhite, born elsewhere	2	2	2	2	6	1
Family history of breast cancer (%)[4]	12	13	12	10	12	13
Benign breast biopsy (%)	16	18	16	12	15	19
Age at menarche (y)	12 (12, 13)	12 (12, 13)	13 (12, 13)	12 (12, 13)	12 (12, 13)	12 (12, 13)
Nulliparous (%)	27	23	24	30	29	25
Age at first full-term pregnancy (y)[5]	26 (23, 29)	26 (23, 29)	26 (23, 29)	27 (24, 30)	26 (23, 29)	25 (23, 28)
Menopausal status and HT use (%)						
Premenopausal	43	32	37	66	43	23
Postmenopausal[6]						
Never used HT	10	14	14	5	12	13
Past HT use	6	8	8	3	7	9
Current use of ET	13	16	15	7	12	19
Current use of EPT	15	17	15	9	13	21
Can't determine menopausal status	8	7	7	8	7	8
Unknown	5	6	5	2	6	6
BMI (kg/m^2)	23.7 (21.3, 27.3)	23.8 (21.5, 27.5)	25.0 (22.0, 29.3)	24.1 (21.5, 28.3)	23.5 (21.2, 27.2)	24.0 (21.7, 27.2)
Height (inches)	65 (63, 67)	65 (63, 67)	65 (63, 67)	65 (63, 67)	65 (63, 66)	65 (63, 67)
Daily caloric intake (kcal/d)	1545 (1242, 1911)	1741 (1430, 2122)	2012 (1687, 2397)	2101 (1797, 2456)	1744 (1409, 2148)	1682 (1386, 2047)
Moderate physical activity (h/wk)	1.7 (0.7, 3.3)	2.0 (1.0, 3.8)	1.7 (0.6, 3.3)	2.0 (0.9, 3.6)	2.0 (0.8, 3.8)	1.9 (0.8, 3.6)
Socioeconomic status (%)						
Deciles 1–5 (low)	21	22	27	22	26	17
Deciles 6–7	26	26	27	28	26	22
Deciles 8–9	35	34	32	35	33	37
Decile 10 (high)	18	19	15	15	16	24

[1] ET, estrogen therapy; EPT, estrogen plus progesterone therapy; HT, hormone therapy.
[2] Women scoring in the highest quintile for more than one pattern are included for each pattern in which they scored high.
[3] Median; IQR in parentheses (all such values).
[4] In a first-degree relative.
[5] Among parous women.
[6] Inludes perimenopausal women.

TABLE 3
Associations between dietary patterns and breast cancer risk[1]

	Quintile					
Dietary pattern	1	2	3	4	5	*P*-trend[2]
Plant-based						
No. of cases	749	807	847	887	850	
RR	1.0	0.97	0.95	0.93	0.85	0.003
95% CI		0.88, 1.07	0.86, 1.05	0.84, 1.04	0.76, 0.95	
High-protein, high-fat						
No. of cases	779	763	849	865	884	
RR	1.0	0.94	1.01	1.02	1.00	0.60
95% CI		0.85, 1.04	0.91, 1.12	0.91, 1.13	0.88, 1.14	
High-carbohydrate						
No. of cases	1,013	906	853	716	652	
RR	1.0	0.93	0.94	0.86	0.91	0.11
95% CI		0.84, 1.02	0.85, 1.04	0.77, 0.97	0.79, 1.05	
Ethnic						
No. of cases	907	854	831	788	760	
RR	1.0	0.99	0.99	0.97	0.94	0.24
95% CI		0.90, 1.09	0.90, 1.09	0.88, 1.07	0.85, 1.05	
Salad and wine						
No. of cases	581	692	856	891	1,120	
RR	1.0	0.95	1.03	0.98	1.12	0.010
95% CI		0.85, 1.06	0.92, 1.14	0.88, 1.09	1.01, 1.25	

[1] Cox proportional hazards regression model with age as the time metric and stratified by age at baseline; adjusted for race-ethnicity/birthplace, family history of breast cancer, age at menarche, parity/age at first full-term pregnancy, average daily caloric intake, physical activity, socioeconomic status, history of a benign breast biopsy and its interaction with time-dependent age, BMI, height, menopausal status/hormone therapy use, and the other 4 dietary patterns.
[2] Likelihood ratio test for trend across dietary pattern quintiles by using an ordinal variable coded as the median value of the quintile.

wine pattern increased the risk of ER+PR+ tumors (RR: 1.29; 95% CI: 1.12, 1.49 for the highest compared with the lowest quintile; *P*-trend < 0.001). Adjustment for alcohol did not affect the risk associated with a plant-based dietary pattern and only slightly attenuated the RR associated with the highest consumption of the salad and wine pattern (RR: 1.26; 95% CI: 1.08, 1.46; *P*-trend = 0.002).

DISCUSSION

Of the 5 major dietary patterns identified in this cohort of California women, only the plant-based and salad and wine patterns were associated with breast cancer risk overall. Greater adherence to the plant-based diet was associated with a reduced risk of breast cancer, whereas greater adherence to the salad and wine pattern was associated with an increased risk. The latter finding was only slightly attenuated when overall alcohol consumption was accounted for.

Dietary patterns that have been characterized by a high vegetable intake have been inversely associated with breast cancer risk in many studies (5, 8, 14, 17, 18, 20, 21, 33). In addition, a meta-analysis of 8 cohort and 8 case-control studies published between 2001 and 2009 found a significantly reduced risk of breast cancer for a prudent/healthy diet (OR: 0.89; 95% CI: 0.82, 0.99 for the highest compared with the lowest categories of intake) (14). Consistent with our finding of a protective effect of a plant-based diet on ER−PR− breast cancer, the Nurses' Health Study found a prudent diet to be associated with a lower risk of ER− breast cancer (21), and the Melbourne Collaborative Cohort Study found that a fruit and salad diet reduced the risk of ER−PR− tumors but not of ER+PR+ tumors (12). In contrast, a reduced risk of invasive breast cancer associated with a healthy Mediterranean diet did not vary by hormone-receptor subtype in the European Prospective Investigation into Cancer and Nutrition cohort (17). Whether hormone levels from endogenous or exogenous sources, as reflected in menopausal status, HT use, and body mass, modify the diet-cancer relation is not clear. Whereas some studies have reported a reduced risk associated with a healthy diet only among postmenopausal women (16, 24), others have found a reduced risk only among premenopausal women (11, 23). The Hormones and Diet in the Etiology of Breast Cancer cohort found that the salad/vegetables dietary pattern was inversely associated with breast cancer risk only among women of normal weight, regardless of menopausal status (5). We observed no statistically significant interactions between the dietary patterns and menopausal status/HT use or BMI. Taken together, our findings and much of the published literature suggest that greater adherence to a healthy diet characterized by greater consumption of plant-based items may reduce the risk of some types of breast cancer.

Many components of a plant-based diet might reduce breast cancer risk. The fiber component of fruit and vegetables, particularly insoluble fiber, may bind with estrogens and lead to excretion, which reduces serum estrogen concentrations (34). A greater intake of soluble fiber, generally associated with a diet characterized by a lower glycemic load, may help maintain lower concentrations of glucose, insulin, and insulin-like growth factors (35, 36). Many fruit and vegetables are also high in antioxidants or contain phytochemicals which have been shown to inhibit the growth of breast cancer cells (37, 38).

Most studies that have used principal component analysis or cluster analysis to evaluate overall diet have identified healthy- and Western-type dietary patterns (6, 11, 14, 15, 39–41). However, other patterns, such as our ethnic or salad and wine patterns, are

TABLE 4
Associations between dietary patterns and breast cancer risk by hormone receptor status[1]

Hormone receptor status and dietary pattern	Quintile 1	2	3	4	5	*P*-trend[2]
ER+PR+ (*n* = 2422 cases)						
Plant-based	1.0	1.00 (0.87, 1.14)	1.00 (0.87, 1.14)	1.01 (0.89, 1.16)	0.91 (0.78, 1.05)	0.19
High-protein, high-fat	1.0	0.94 (0.82, 1.07)	1.03 (0.90, 1.18)	1.05 (0.91, 1.20)	1.02 (0.86, 1.21)	0.52
High-carbohydrate	1.0	0.95 (0.84, 1.08)	0.96 (0.84, 1.10)	0.89 (0.76, 1.03)	0.91 (0.76, 1.10)	0.26
Ethnic	1.0	0.98 (0.86, 1.10)	0.99 (0.88, 1.12)	0.93 (0.82, 1.06)	0.89 (0.78, 1.02)	0.07
Salad and wine	1.0	1.02 (0.88, 1.18)	1.15 (0.99, 1.32)	1.05 (0.91, 1.21)	1.29 (1.12, 1.49)	<0.001
ER+PR− (*n* = 509 cases)						
Plant-based	1.0	1.08 (0.81, 1.46)	1.08 (0.80, 1.45)	1.05 (0.77, 1.42)	1.03 (0.74, 1.41)	0.94
High-protein, high-fat	1.0	0.95 (0.72, 1.26)	0.94 (0.70, 1.26)	1.01 (0.74, 1.37)	0.90 (0.62, 1.31)	0.70
High-carbohydrate	1.0	0.84 (0.65, 1.08)	0.85 (0.64, 1.13)	0.67 (0.48, 0.94)	0.69 (0.45, 1.06)	0.05
Ethnic	1.0	1.07 (0.81, 1.42)	1.08 (0.81, 1.43)	1.39 (1.06, 1.83)	1.03 (0.76, 1.40)	0.45
Salad and wine	1.0	0.85 (0.62, 1.19)	0.96 (0.70, 1.31)	1.03 (0.75, 1.40)	1.10 (0.80, 1.50)	0.22
ER−PR− (*n* = 514 cases)						
Plant-based	1.0	0.84 (0.64, 1.11)	0.84 (0.63, 1.10)	0.86 (0.65, 1.15)	0.66 (0.48, 0.91)	0.03
High-protein, high-fat	1.0	0.91 (0.68, 1.22)	1.01 (0.76, 1.35)	1.11 (0.82, 1.50)	1.02 (0.70, 1.47)	0.64
High-carbohydrate	1.0	0.85 (0.65, 1.11)	0.79 (0.58, 1.06)	0.97 (0.70, 1.33)	0.99 (0.66, 1.48)	0.88
Ethnic	1.0	1.14 (0.88, 1.49)	0.92 (0.70, 1.22)	0.90 (0.68, 1.20)	1.06 (0.79, 1.42)	0.89
Salad and wine	1.0	0.81 (0.61, 1.09)	0.75 (0.56, 1.00)	0.78 (0.58, 1.05)	0.85 (0.63, 1.14)	0.41

[1] Values are RRs; 95% CIs in parentheses. Cox proportional hazards regression models with age as the time metric and stratified by age at baseline; adjusted for race-ethnicity/birthplace, family history of breast cancer, age at menarche, parity/age at first full-term pregnancy, average daily caloric intake, physical activity, socioeconomic status, history of a benign breast biopsy and its interaction with time-dependent age, BMI, height, menopausal status and hormone therapy use, and the other 4 dietary patterns. ER–, estrogen receptor negative; ER+, estrogen receptor positive; PR–, progesterone receptor negative; PR+, progesterone receptor positive.

[2] Likelihood ratio test for trend across dietary pattern quintiles by using an ordinal variable coded as the median value of the quintile.

observed to be prominent in a few populations (12, 13, 39, 40, 42). These studies thus allow for the examination of a greater diversity of dietary intake. Some (6, 13, 17, 22), but not all (5, 18, 19), studies that evaluated dietary patterns that have identified individual patterns, including high factor loadings for alcoholic beverages, have found a drinker pattern associated with increased breast cancer risk. Furthermore, similar to our study, the E3N- European Prospective Investigation into Cancer and Nutrition study found a significantly increased risk with an alcohol/Western diet for ER+PR+, but not for ER−PR−, tumors (17). A meta-analysis that included 4 of the above studies (6, 18, 19, 22) found a significantly increased risk of breast cancer associated with the drinker dietary pattern (OR: 1.21; 95% CI: 1.04, 1.41 for the highest compared with the lowest category) (14). Consistent with these studies, the salad and wine pattern in our study was associated with an increased risk of ER+PR+ breast cancer—an association that was not accounted for by alcohol consumption. Furthermore, given the high reproducibility (r = 0.87) and validity ($r = \sim 0.75$) for alcohol intake in our validation study (26), it is unlikely that this association was due to residual confounding from misclassification resulting from errors in self-reported alcohol intake. However, because of the synergistic effects between alcohol intake and HT use on breast cancer risk (30–32, 43–47) and the higher prevalence of HT use among those who consumed the salad and wine pattern, the influence of HT use on this association warrants further investigation.

A major strength of this analysis was its basis in a large diverse cohort with dietary intake data collected before breast cancer diagnosis and based on a widely used and validated food-frequency questionnaire (26). Also, the reporting of cancer outcomes is essentially complete for cohort members who reside in California. Several limitations should also be noted. The dietary intake data for these analyses were based on the 1-y period preceding the baseline assessment. To the extent that diet has changed or it is diet during other potentially critical periods of life, such as puberty, that are important, these analyses become less meaningful. In addition, several studies have evaluated raw vegetable intakes and found them to be inversely related to breast cancer risk (48). However, the CTS dietary assessment did not separate cooked from raw vegetables; therefore, if a separate dietary pattern characterized by raw, as opposed to cooked, vegetable intake were present, we would not have been able to distinguish it here.

Finally, there are limitations to using PCFA. Dietary patterns identified with PCFA may not have distinct biological effects on the body; thus, their relation with health or disease risk may be attenuated (49, 50). Also, the dietary patterns in this study explained only 19% of the variance in dietary intake; while typical (49, 50), it is still rather low. Whereas the interpretation for those falling into the lowest and highest quintiles for a dietary pattern was likely to represent a clear distinction in the consumption of the foods that define that pattern, the interpretation in the middle quintiles is less clear. PCFA also has been criticized because it captures dietary patterns that are relatively unique to specific populations (51). However, because the literature on dietary patterns is growing, similar core patterns appear to be present in most populations, with specific patterns evident in different populations. Exploring dietary patterns from diverse populations may help identify combinations of foods that decrease the risk of specific diseases, such as has been observed for heart disease and the Mediterranean diet (52). Finally, despite its limitations, the PCFA approach to studying dietary intake reflects the combinations of foods that are consumed and the nutrient interactions that may thus occur.

In conclusion, our study found that a plant-based diet was associated with a reduced risk of ER−PR− breast cancer. In addition, despite the inclusion of healthy foods in the salad and wine dietary pattern, the risk of ER+PR+ breast cancer remained elevated among women consuming this type of diet, although the effect of HT use on this association warrants further investigation.

We thank the members of the CTS Steering Committee, who are responsible for the formation and maintenance of the cohort within which this study was conducted but who are not included as authors.

The authors' responsibilities were as follows—LBL and PLH-R: designed the specific research project and took primary responsibility for the final content of the manuscript; LBL, CAC, DOS, GU, and PLH-R: contributed to the formation and maintenance of the cohort, including the data collection; LBL and AJC: performed the data analysis; and LBL, AJC, and PLH-R: took primary responsibility for writing the manuscript. All authors made intellectual contributions to and read and approved the final manuscript. CAC served as an expert witness for plaintiffs in litigation regarding HT and breast cancer. None of the other authors had any conflicts of interest related to this study. The funding sources did not contribute to the design or conduct of the study or to the writing or submission of the manuscript.

REFERENCES

1. Willett WC. Diet and breast cancer. J Intern Med 2001;249:395–411.
2. World Cancer Research Fund/American Institute for Cancer Research. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Washington, DC: AICR, 2007.
3. Michels KB, Mohllajee AP, Roset-Bahmanyar E, Beehler GP, Moysich KB. Diet and breast cancer: a review of the prospective observational studies. Cancer 2007;109(suppl):2712–49.
4. Cappellani A, Di Vita M, Zanghi A, Cavallaro A, Piccolo G, Veroux M, Berretta M, Malaguarnera M, Canzonieri V, Lo Menzo E. Diet, obesity and breast cancer: an update. Front Biosci (Schol Ed) 2012;4:90–108.
5. Sieri S, Krogh V, Pala V, Muti P, Micheli A, Evangelista A, Tagliabue G, Berrino F. Dietary patterns and risk of breast cancer in the ORDET cohort. Cancer Epidemiol Biomarkers Prev 2004;13:567–72.
6. Terry P, Suzuki R, Hu FB, Wolk A. A prospective study of major dietary patterns and the risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2001;10:1281–5.
7. Cui X, Dai Q, Tseng M, Shu XO, Gao YT, Zheng W. Dietary patterns and breast cancer risk in the Shanghai breast cancer study. Cancer Epidemiol Biomarkers Prev 2007;16:1443–8.
8. Hirose K, Matsuo K, Iwata H, Tajima K. Dietary patterns and the risk of breast cancer in Japanese women. Cancer Sci 2007;98:1431–8.
9. Mannistö S, Dixon LB, Balder HF, Virtanen MJ, Krogh V, Khani BR, Berrino F, van den Brandt PA, Hartman AM, Pietinen P, et al. Dietary patterns and breast cancer risk: results from three cohort studies in the DIETSCAN project. Cancer Causes Control 2005;16:725–33.
10. Adebamowo CA, Hu FB, Cho E, Spiegelman D, Holmes MD, Willett WC. Dietary patterns and the risk of breast cancer. Ann Epidemiol 2005;15:789–95.
11. Agurs-Collins T, Rosenberg L, Makambi K, Palmer JR, Adams-Campbell L. Dietary patterns and breast cancer risk in women participating in the Black Women's Health Study. Am J Clin Nutr 2009; 90:621–8.
12. Baglietto L, Krishnan K, Severi G, Hodge A, Brinkman M, English DR, McLean C, Hopper JL, Giles GG. Dietary patterns and risk of breast cancer. Br J Cancer 2011;104:524–31.
13. Bessaoud F, Tretarre B, Daures JP, Gerber M. Identification of dietary patterns using two statistical approaches and their association with breast cancer risk: a case-control study in southern france. Ann Epidemiol 2012;22:499–510.
14. Brennan SF, Cantwell MM, Cardwell CR, Velentzis LS, Woodside JV. Dietary patterns and breast cancer risk: a systematic review and meta-analysis. Am J Clin Nutr 2010;91:1294–302.
15. Buck K, Vrieling A, Flesch-Janys D, Chang-Claude J. Dietary patterns and the risk of postmenopausal breast cancer in a German case-control study. Cancer Causes Control 2011;22:273–82.
16. Butler LM, Wu AH, Wang R, Koh WP, Yuan JM, Yu MC. A vegetable-fruit-soy dietary pattern protects against breast cancer among postmenopausal Singapore Chinese women. Am J Clin Nutr 2010;91:1013–9.
17. Cottet V, Touvier M, Fournier A, Touillaud MS, Lafay L, Clavel-Chapelon F, Boutron-Ruault MC. Postmenopausal breast cancer risk and dietary patterns in the E3N-EPIC prospective cohort study. Am J Epidemiol 2009;170:1257–67.
18. Ronco AL, De Stefani E, Boffetta P, Deneo-Pellegrini H, Acosta G, Mendilaharsu M. Food patterns and risk of breast cancer: a factor analysis study in Uruguay. Int J Cancer 2006;119:1672–8.
19. Nkondjock A, Ghadirian P. Associated nutritional risk of breast and colon cancers: a population-based case-control study in Montreal, Canada. Cancer Lett 2005;223:85–91.
20. Velie EM, Schairer C, Flood A, He JP, Khattree R, Schatzkin A. Empirically derived dietary patterns and risk of postmenopausal breast cancer in a large prospective cohort study. Am J Clin Nutr 2005;82: 1308–19.
21. Fung TT, Hu FB, Holmes MD, Rosner BA, Hunter DJ, Colditz GA, Willett WC. Dietary patterns and the risk of postmenopausal breast cancer. Int J Cancer 2005;116:116–21.
22. De Stefani E, Deneo-Pellegrini H, Boffetta P, Ronco AL, Aune D, Acosta G, Mendilaharsu M, Brennan P, Ferro G. Dietary patterns and risk of cancer: a factor analysis in Uruguay. Int J Cancer 2009;124: 1391–7.
23. Cade JE, Taylor EF, Burley VJ, Greenwood DC. Does the Mediterranean dietary pattern or the Healthy Diet Index influence the risk of breast cancer in a large British cohort of women? Eur J Clin Nutr 2011; 65:920–8.
24. Trichopoulou A, Bamia C, Lagiou P, Trichopoulos D. Conformity to traditional Mediterranean diet and breast cancer risk in the Greek EPIC (European Prospective Investigation into Cancer and Nutrition) cohort. Am J Clin Nutr 2010;92:620–5.
25. Bernstein L, Anton-Culver H, Deapen D, Horn-Ross PL, Peel D, Reynolds P, Ross RK, Sullivan-Halley J, West DW, Wright W, et al. High breast cancer rates among California teachers: Results from the California Teachers Study Cohort. Cancer Causes Control 2002;13:625–35.
26. Horn-Ross PL, Lee VS, Collins CN, Stewart SL, Canchola AJ, Lee MM, Reynolds P, Clarke CA, Bernstein L, Stram DO. Dietary assessment in the California Teachers Study: reproducibility and validity. Cancer Causes Control 2008;19:595–603.
27. Reynolds P, Hurley S, Goldberg DE, Anton-Culver H, Bernstein L, Deapen D, Horn-Ross PL, Peel D, Pinder R, Ross RK, et al. Regional variations in breast cancer among California teachers. Epidemiology 2004;15:746–54.
28. Chang ET, Lee VS, Canchola AJ, Dalvi TB, Clarke CA, Reynolds P, Purdie DM, Stram DO, West DW, Ziogas A, et al. Dietary patterns and risk of ovarian cancer in the California Teachers Study cohort. Nutr Cancer 2008;60:285–91.
29. Hatcher L. A step-by-step approach to using the SAS System for factor analysis and structural equation modeling. Cary, NC: SAS Institute, Inc, 1994.
30. Horn-Ross PL, Canchola AJ, West DW, Stewart SL, Bernstein L, Deapen D, Pinder R, Ross RK, Anton-Culver H, Peel D, et al. Patterns of alcohol consumption and breast cancer risk in the California Teachers Study cohort. Cancer Epidemiol Biomarkers Prev 2004;13:405–11.
31. Saxena T, Lee E, Henderson KD, Clarke CA, West D, Marshall SF, Deapen D, Bernstein L, Ursin G. Menopausal hormone therapy and subsequent risk of specific invasive breast cancer subtypes in the California Teachers Study. Cancer Epidemiol Biomarkers Prev 2010; 19:2366–78.
32. Marshall SF, Clarke CA, Deapen D, Henderson KD, Largent J, Neuhausen SL, Reynolds P, Ursin G, Horn-Ross PL, Stram DO, et al. Recent breast cancer incidence trends according to hormone therapy use: the California Teachers Study. Breast Cancer Res 2010;12:R4.
33. Zhang CX, Ho SC, Fu JH, Cheng SZ, Chen YM, Lin FY. Dietary patterns and breast cancer risk among Chinese women. Cancer Causes Control 2011;22:115–24.
34. Rose DP, Goldman M, Connolly JM, Strong LE. High-fiber diet reduces serum estrogen concentrations in premenopausal women. Am J Clin Nutr 1991;54:520–5.
35. Key TJ, Appleby PN, Reeves GK, Roddam AW. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data of 17 prospective studies. Lancet Oncol 2010;11:530–42.
36. Lawlor DA, Smith GD, Ebrahim S. Hyperinsulinaemia and increased risk of breast cancer: findings from the British Women's Heart and Health Study. Cancer Causes Control 2004;15:267–75.

37. Reagan-Shaw S, Eggert D, Mukhtar H, Ahmad N. Antiproliferative effects of apple peel extract against cancer cells. Nutr Cancer 2010;62: 517–24.
38. Jo EH, Kim SH, Ahn NS, Park IS, Hwang IW, Lee YS, Kang KS. Efficacy of sulforaphane is mediated by p38 MAP kinase and caspase-7 activation in ER-positive and COX-2-expressed human breast cancer cells. Eur J Cancer Prev 2007;16:505–10.
39. Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. Am J Epidemiol 1998;148:4–16.
40. Engeset D, Alsaker E, Ciampi A, Lund E. Dietary patterns and lifestyle factors in the Norwegian EPIC cohort: the Norwegian Women and Cancer (NOWAC) study. Eur J Clin Nutr 2005;59:675–84.
41. Huijbregts PP, Feskens EJ, Kromhout D. Dietary patterns and cardiovascular risk factors in elderly men: the Zutphen Elderly Study. Int J Epidemiol 1995;24:313–20.
42. Jordan I, Hebestreit A, Swai B, Krawinkel MB. Dietary patterns and breast cancer risk among women in northern Tanzania: a case-control study. Eur J Nutr 2012;52:905–15.
43. Horn-Ross PL, Canchola AJ, Bernstein L, Clarke CA, Lacey JV, Neuhausen SL, Reynolds P, Ursin G. Alcohol consumption and breast cancer risk among postmenopausal women following the cessation of hormone therapy use: the California Teachers Study. Cancer Epidemiol Biomarkers Prev 2012;21:2006–13.
44. Lew JQ, Freedman ND, Leitzmann MF, Brinton LA, Hoover RN, Hollenbeck AR, Schatzkin A, Park Y. Alcohol and risk of breast cancer by histologic type and hormone receptor status in postmenopausal women: the NIH-AARP Diet and Health Study. Am J Epidemiol 2009; 170:308–17.
45. Nielsen NR, Gronbaek M. Interactions between intakes of alcohol and postmenopausal hormones on the risk of breast cancer. Int J Cancer 2008;122:1109–13.
46. Zhang SM, Lee I-M, Manson JE, Cook NR, Willett WC, Buring JE. Alcohol consumption and breast cancer risk in the Women's Health Study. Am J Epidemiol 2007;165:667–76.
47. Smith-Warner SA, Spiegelman D, Yaun SS, van den Brandt PA, Folsom AR, Goldbohm RA, Graham S, Holmberg L, Howe GR, Marshall JR, et al. Alcohol and breast cancer in women: a pooled analysis of cohort studies. JAMA 1998;279:535–40.
48. Link LB, Potter JD. Raw versus cooked vegetables and cancer risk. Cancer Epidemiol Biomarkers Prev 2004;13:1422–35.
49. Edefonti V, Randi G, Decarli A, La Vecchia C, Bosetti C, Franceschi S, Dal Maso L, Ferraroni M. Clustering dietary habits and the risk of breast and ovarian cancers. Ann Oncol 2009;20:581–90.
50. Edefonti V, Randi G, La Vecchia C, Ferraroni M, Decarli A. Dietary patterns and breast cancer: a review with focus on methodological issues. Nutr Rev 2009;67:297–314.
51. Jacques PF, Tucker KL. Are dietary patterns useful for understanding the role of diet in chronic disease? Am J Clin Nutr 2001;73:1–2.
52. Willett WC. The Mediterranean diet: science and practice. Public Health Nutr 2006;9(1A):105–10.