◼ ORIGINAL ARTICLE

Cohort Study of Risk Factors for Breast Cancer in Post Menopausal Women

Arthur J. Hartz, Tao He

Health Services Research, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT, USA

OBJECTIVES: The present study assessed more than 800 potential risk factors to identify new predictors of breast cancer and compare the independence and relative importance of established risk factors.

METHODS: Data were collected by the Women's Health Initiative and included 147,202 women ages 50 to 79 who were enrolled from 1993 to 1998 and followed for 8 years. Analyses performed in 2011 and 2012 used the Cox proportional hazard regression to test the association between more than 800 baseline risk factors and incident breast cancer.

RESULTS: Baseline factors independently associated with subsequent breast cancer at the $p < 0.001$ level (in decreasing order of statistical significance) were breast aspiration, family history, age, weight, history of breast biopsies, estrogen and progestin use, fewer live births, greater age at menopause, history of thyroid cancer, breast tenderness, digitalis use, alcohol intake, white race, not restless, no vaginal dryness, relative with prostate cancer, colon polyps, smoking, no breast augmentation, and no osteoporosis. Risk factors previously reported that were not independently associated with breast cancer in the present study included socioeconomic status, months of breast feeding, age at first birth, adiposity measures, adult weight gain, timing of initiation of hormone therapy, and several dietary, psychological, and exercise variables. Family history was not found to alter the risk associated with other factors.

CONCLUSIONS: These results suggest that some risk factors not commonly studied may be important for breast cancer and some frequently cited risk factors may be relatively unimportant or secondary.

KEY WORDS: Breast cancer, Risk factors, Obesity, Postmenopausal hormone replacement therapy, Thyroid cancer, Breast augmentation

INTRODUCTION

Breast cancer risk factors have been studied often. The present study has the advantage of simultaneously considering many risk factors. This makes it possible to find new risk factors, assess the relative importance of risk factors, and assess which risk factors are secondary (their association is removed by adjusting for other risk factors). The data for the present study were collected by the Women's Health Initiative (WHI). Because the WHI is a prospective study with a large, comprehensive, meticulously collected database, it is especially well-suited for evaluating risk factors.

MATERIALS AND METHODS

The WHI study design has been described in detail [1]. In brief, it was a long-term national health study that focused on strategies for preventing heart disease, breast and colorectal cancer, and osteoporosis in postmenopausal women. Women between the ages of 50 and 79 were enrolled in either an ob-

Correspondence: **Arthur J. Hartz, MD, PhD**
Health Services Research, Huntsman Cancer Institute, University of Utah School of Medicine, 7450 Washington Ave., University City, MO 63130, USA
Tel: +1-314-769-9247, Fax: +1-314-362-4247
E-mail: hartzarthur@gmail.com

servational study (OS) or a randomized controlled trial (RCT) from 1993 to 1998 at 40 clinical centers throughout the United States and followed for a median time of 8 years to ascertain outcomes. The study was approved by institutional review boards, and all participants signed informed consent forms.

Participants available for analysis included 161,807 WHI participants: 93,675 from the OS, 16,608 from the RCT of estrogen plus progesterone (E+P), 10,739 from the RCT of estrogen only (estrogen-alone), and 40,785 additional women who were in the diet study and not in a RCT of hormone therapy. To minimize the inclusion of participants with a history of breast cancer, we excluded from the present analysis those who reported a history of breast cancer, a breast removed, or current antineoplastic therapy. Of the 161,807 participants in the dataset we excluded 7,906 because of a possible history of breast cancer and 825 patients who could not influence the analysis because they did not have follow-up. An additional 5,874 participants were excluded from the final analyses because they did not have data on all factors independently associated at the $p < 0.001$ level with the development of breast cancer. There were 147,202 women included in the final analyses.

Data

Several data sources were used to determine breast cancer recurrence and the date of this recurrence [1]. For follow-up and outcome ascertainment all subjects completed a self-administered, self-report. This report was completed semiannually by the RCT participants and annually by the OS participants. Medical records were reviewed for patients who died in the hospital, and autopsy reports were reviewed for patients who had an autopsy. Only the death certificates were reviewed for patients who died outside of the hospital without an autopsy.

There were 869 factors that characterized subjects at baseline that were evaluated for an association with the risk of breast cancer for all participants and an additional 102 factors were evaluated for an association with participants in the OS. Types of information included demographic, general health, clinical and anthropometric, functional status, healthcare behaviors, reproductive, medical history, family history, personal habits, thoughts and feelings, therapeutic class of medication, hormones, supplements, and dietary intake. Scales used in this study that we anticipated might be important included stressful life events (a scale from 0 to 33 from the Alameda County Study), optimism (a 6 to 30 scale from the Life Orientation Test-Revised) and depression (a 0 to 1 scale from the Center for Epidemiological Studies).

Statistical methods

Analyses were performed in 2011 and 2012 using the Cox proportional hazard regression model to find the hazard ratio of breast cancer associated with a factor after adjusting for other risk factors. We first tested the statistical significance of each potential risk factor after adjusting for only the study that recruited the patient, age, and race. All variables that were statistically significant at the $p < 0.05$ level were then included in a backwards stepwise Cox proportional hazard regression analysis, and variables that were statistically significant at the $p < 0.001$ level were retained in the model. We then retested the statistical significance of all variables not in the model to determine if any should be added because they would be significant at the $p < 0.001$ level. The significance level of the variables was not adjusted for multiple comparisons because this might obscure meaningful results.

The hazard ratio was usually reported for a one unit increase in the risk factor. For age risk was measured for a 10 year increase, and where noted the hazard ratio was for an increase in one standard deviation of the risk factor. For categorical variables a reference category was chosen, and the hazard ratios were shown for other categories compared to the reference category. Some ordinal variables were also divided into categories, and in addition to an overall hazard ratio and significance test for the ordinal variable there was a hazard ratio comparing a given category to the reference category.

To test whether the risks associated with some variables were influenced by others, we tested pairs of the statistically significant independent variables for interaction.

In addition to testing the variables in the complete dataset, we also tested the results for participants in each of three subsets: 1) the OS, 2) the RCT of diet, and 3) the combination of the two RCTs of hormone therapy that were each relatively small.

Missing data for variables that were not independently significant were imputed prior to the analysis of these variables by the mean value for ordinal or binary variables and the mode value for categorical variables.

Statistical analyses were performed using SAS version 9.0 (SAS Inc., Cary, NC, USA).

RESULTS

The demographic characteristics of the participants are shown in Table 1. Most were white, between the ages of 55 and 70, and with more than a high school education. The regions of the United States were equally represented.

Demographic factors

As shown in Table 2 three demographic factors (age, white race, and specific study to which subjects were enrolled) have an independent statistically significant association ($p < 0.001$)

with breast cancer after adjusting for all factors independently significant at the $p < 0.001$ level. The association of age and race with breast cancer did not vary significantly across the three

Table 1. Demographic characteristics

Characteristics	Categories	% (n = 147,202)
Age (yr)	49-55	33.60
	56-69	44.94
	70-81	21.45
Race	Caucasian	82.91
	Non-Caucasian	17.09
Where live in US	Northeast	23.18
	South	25.43
	Midwest	21.96
	West	29.43
Education level	High school graduate or less	23.00
	Post high school	37.75
	College graduate or higher	39.24
Income (USD)	Unknown	6.57
	< 35,000	38.12
	35,000-75,000	37.88
	> 75,000	17.42
Insurance type (can have > 1)	Pre-paid private insurance	40.26
	Other private insurance	44.77
	Medicare	36.22
	Medicaid	1.33
	Other	6.84
	None	4.56
WHI data set	Observational study	56.06
	RCT estrogen alone	6.83
	RCT estrogen plus progesterone	10.75
	RCT diet (not in other RCTs)	26.36

USD, United States dollar; WHI, Women's Health Initiative; RCT, randomized controlled trial.

datasets. Other variables in Table 2 (college education, income, professional or managerial occupation, working with hair dyes for one year) had an association with breast cancer at the $p < 0.0001$ level when only adjusting for age, race, and specific study. However, the significance was reduced after adjusting for the other variables that were statistically significant at the $p < 0.001$ level. Working with hair dyes seemed to be associated with breast cancer in a univariate analysis only because it was a marker for socioeconomic status, 3.5% of 34,234 women with family incomes of less than $50,000 reported working with hair dyes as compared to 1.2% of 35,094 women with family incomes of $50,000 or more.

Breast and reproductive factors

The highest chi-squared values in Table 3 were for the number of needle aspirations of breast cysts, having a first degree relative with breast cancer, and having a biopsy of lesions suspicious of breast carcinoma. A history of hormone therapy was also important. The number of years of using hormone therapy and whether or not hormone therapy was used at baseline were independent risk factors. The risk of breast cancer was 55% greater for women on estrogen and progesterone for more than 15 years compared with women not on hormones. The hazard ratio for women taking estrogen plus progesterone at baseline, 1.29, was greater than the hazard ratio for women taking estrogen alone, 1.14.

The best way to show how hazard ratios for continuous variables varied across data sets was to compare the hazard ratio for the continuous form of variable. The confidence intervals for categories of continuous variables in the individual datasets

Table 2. The association of demographic factors with breast cancer

Factor (scale)	% or (mean)	All subjects			OS	RCT diet	RCT HT
		HR	95% CI	χ^2	HR	HR	HR
Independently significant[†]							
Age*	(63.1)	1.22	1.17-1.26	102.9[‖]	1.24[‖]	1.14[‖]	1.25[‖]
Study[‡]							
OS	56.1	1.00			NA	NA	NA
RCT E+P	10.7	0.91	0.83-1.00	3.8	NA	NA	NA
RCT E alone	6.8	0.83	0.73-0.94	9.0[‖]	NA	NA	NA
RCT diet	26.3	1.14	1.07-1.22	15.0[‖]	NA	NA	NA
White race	82.8	1.18	1.10-1.28	18.0[‖]	1.17[‖]	1.24[‖]	1.13
Not independently significant[†]							
College education	39.2	1.09	1.03-1.15	10.2[‖]	1.09	1.14	0.99
Income* (1-8)	(4.2)	1.03	1.01-1.06	5.6	1.02	1.07[‖]	1.02
Professional or managerial	38.7	1.03	0.98-1.08	1.1	1.03	1.11	1.00
Job with hair dyes[§]	2.3	NA	NA	NA	0.67[‖]	NA	NA

HR, hazard ratio; CI, confidence interval; OS, observational study; RCT, randomized controlled trial; RCT HT, combination of the RCTs for E+P and for E-alone; E+P, estrogen plus progesterone; E-alone, estrogen alone.
*The hazard ratio was computed for an increase in the variable of one standard deviation; [†]A variable is independently significant if $p < 0.001$ after adjusting for all other independently significant variables; [‡]Hazard ratios for a given study compare the breast cancer risk for participants in that study to the participants in the observational study; [§]Subjects who worked with hair dyes for > 1 year were compared to all other subjects; [‖]The hazard ratio is statistically significant at $p < 0.01$.

Table 3. The association* of familial, breast, and reproductive factors with breast cancer

Factor	All subjects				Individual dataset		
	% or (mean)	HR*	95% CI	χ^2	OS HR*	Diet RCT HR*	HT RCT HR*
Independently significant†							
Br aspirations		1.18	1.15-1.22	123.9	1.15[II]	1.23[II]	1.30[II]
0	86.8	1.00					
1-2	10.7	1.36	1.26-1.46	62.5			
3-4	2.5	1.85	1.64-2.08	99.9			
Br Ca in 1st degree relative	13.7	1.41	1.32-1.51	107.6	1.38[II]	1.42[II]	1.55[II]
Br biopsy	24.9	1.33	1.26-1,42	86.9	1.38[II]	1.34[II]	1.10
Years HT‡		1.11	1.08-1.14	59.7	1.10[II]	1.13[II]	1.01
None	73.6	1.00					
< 5 yr	13.4	1.11	1.02-1.20	5.5		-	
5-15 yr	11.2	1.32	1.21-1.44	38.9			
> 15 yr	1.92	1.56	1.34-1.82	33.6			
Baseline E+P	17.2	1.29	1.19-1.41	35.4	1.33[II]	NA	1.30[II]
Baseline E-alone	17.8	1.14	1.05-1.23	9.90	1.20[II]	NA	0.79
Live births		0.95	0.93-0.96	52.0	0.94[II]	0.94[II]	0.96
0	12.5	1.00					
1-4	73.6	0.86	0.80-0.93	15.5			
5-6	10.6	0.72	0.64-0.80	35.5			
>7	3.4	0.61	0.51-0.73	28.4			
Age menopause (HR for 10 yr)	(48.1)	1.11	1.07-1,15	26.5	1.12[II]	1.12[II]	1.03
Br disease§	30.4	NA	NA	NA	1.19[II]	NA	NA
Br tenderness	18.9	1.13	1.08-1.18	24.1	1.12[II]	1.15[II]	1.05
Vaginal dryness	27.1	0.93	0.89-0.97	12.3	0.95	0.91	0.91
Br implants	1.2	0.59	0.43-0.81	11.0	0.64[II]	0.39	0.86
Not independently significant†							
Mammogram ever	96.4	1.24	1.05-1.48	6.0	1.11	1.45	1.34
1st birth > 30 yr	28.1	1.05	0.99-1.12	2.4	1.03	1.05	1.12
Oophorectomy < 40 yr	4.3	0.84	0.72-0.98	5.1	0.89	0.90	0.50[II]
Age menarche‡	12-13 (mn)	0.98	0.96-1.00	6.5	0.98	0.99	0.95
Breastfed ≥ 24 mo	5.5	0.90	0.80-1.02	2.6	1.01	0.74	0.84

Br, breast; CI, confidence interval; E+P, estrogen plus progesterone at baseline; E only, estrogen alone at baseline; HR, hazard ratio; HT, either E+P or E-alone; Mn, mean; OS, observational study; RCT, randomized controlled trial.

*Hazard ratios were computed for an increase of 1 unit of the ordinal variable or comparison to the reference category of that variable. This reference category has a hazard ratio of 1.00. Hazard ratios were only compared for ordinal variables; other comparisons would be based on less precise estimates; †Factors are labeled as independently significant if their p-value < 0.001 after adjusting for all other independently significant variables. Listed factors that were not independently significant had a p value of < 0.001 after adjusting for age, race and study. Their tabulated chi-squared value were obtained after adjusting for the independently significant variables; ‡The hazard ratio was computed for an increase in the variable of one standard deviation; §The doctor said there was benign breast disease. This information was only collected for patients in the observational study. The hazard ratio for benign breast disease was adjusted for all other variables with p < 0.001 including breast aspiration and breast biopsy; [II]The hazard ratio is statistically significant at p < 0.01 in this dataset.

were wide and comparisons were less meaningful. The number of years of hormone use had a stronger association with subsequent breast cancer in the observational and RCT for diet datasets than it did in the RCTs for hormone therapy in these RCTs the past use of hormones had no association with the future use, which was determined by random assignment. The hazard ratios were similar for the observational study and the RCT for baseline use of estrogen plus progesterone but not for baseline use of estrogen alone.

The risk of breast cancer decreased for women as they had more children, up to a 39% reduction for women who had seven or more births. Later age at menopause, breast tenderness and benign breast disease were associated with a higher risk, while vaginal dryness and breast augmentation were associated with lower risk.

Participants who had a mammogram were at higher risk for breast cancer although this was greatly reduced after taking into account the variables that were independently significant. There may have been a protective effect of having an oophorectomy prior to age 40. Neither first birth after age 30 nor breast feeding for at least 2 years had a statistically significant independent association with breast cancer. Later age at menarche had a protective but weak association with breast cancer.

Health and health behaviors

As shown in Table 4 the factor in this group most significantly associated with breast cancer was weight (chi-squared = 93.8). The chi-squared value for weight was higher than for waist measurement, body mass index, or waist-hip ratio. After adjusting for weight, the chi-squared values for these variables were not

Table 4. The association* of health and behaviors with the development of breast cancer

Factor	All subjects				Individual dataset		
	%	HR*	95% CI	χ^2	OS HR*	Diet RCT HR*	HT RCT HR*
Independently significant[†]							
Weight at baseline[‡]		1.14	1.11-1.17	93.8	1.14[‖]	1.12[‖]	1.19[‖]
< 150	43.5	1.00					
150-200	13.2	1.21	1.15-1.28	46.8			
> 200	43.4	1.43	1.32-1.55	74.5			
History of thyroid cancer	0.4	2.04	1.55-2.68	26.3	2.09[‖]	2.56[‖]	0.42[‖]
Digitalis use	1.8	1.46	1.24-1.72	21.0	1.58[‖]	1.13	1.61
Alcohol servings per wk[‡]		1.06	1.03-1.09	18.2	1.11[‖]	0.99	1.01
None		1.00					
< 1/mo		1.04	0.95-1.14	0.73			
< 1/wk		1.06	0.98-1.14	2.01			
> 1/wk		1.13	1.05-1.20	12.4			
Restless	34.8	0.92	0.88-0.96	15.2	0.90[‖]	0.94	0.93
Relative with prostrate cancer	9.4	1.15	1.06-1.25	11.7	1.19[‖]	1.20	0.92
Colon polyps	8.6	1.16	1.06-1.26	11.2	1.09	1.22	1.34[‖]
Cigarettes/d[‡] (1-7)		1.04	1.02-1.07	11.0	1.05[‖]	1.03	1.04
None	55.6	1.00					
< 15/d	23.4	1.05	0.98-1.12	2.0			
≥ 15/d	17.1	1.12	1.05-1.19	11.6			
Osteoporosis ever	7.3	0.84	0.76-0.93	10.5	0.79[‖]	0.92	0.99
Not independently significant[†]							
Change from min adult weight[‡,§]	NA	NA	NA	NA	1.05	NA	NA
Trouble sleeping[‡]	41.3	0.98	0.96-1.00	3.1	0.99	1.17[‖]	0.92

HR, hazard ratio; CI, confidence interval; OS, observational study; RCT, randomized controlled trial; RCT HT, a combination of two RCTs; Min, minimum.
*Hazard ratios were computed for an increase of 1 unit of the ordinal variable or comparison to the reference category of that variable. This reference category has a hazard ratio of 1.00. Hazard ratios were only compared for ordinal variables; other comparisons would be based on less precise estimates; [†]Factors are labeled as independently significant if their p-value < 0.001 after adjusting for all other independently significant variables. Listed factors that were not independently significant had a p value of < 0.001 after adjusting for age, race and study. Their tabulated chi-squared value were obtained after adjusting for the independently significant variables. [‡]Change from minimum adult weight was only collected for participants in the observational study; [§]The hazard ratio was computed for an increase in the variable of one standard deviation; [‖]The hazard ratio is statistically significant at p < 0.01 in this dataset.

statistically significant. It was only possible to calculate change in weight (i.e., the difference between current weight and minimum adult weight) in the observational study. Change in weight was not significant after adjusting for current weight although it was highly significant prior to this adjustment.

The risk of breast cancer also increased with history of thyroid cancer, the use of cardiotonic medications (which additional review of the data showed was almost always digitalis), greater alcohol consumption, a history of colon polyps, greater smoking, and a relative with prostate cancer. It decreased for women who reported themselves to be restless and fidgety or having a history of osteoporosis. Although the hazard ratios for some risk factors (e.g., thyroid cancer and relatives with prostate cancer) varied considerably among the datasets, this variation was not statistically significant.

Nonsignificant factors hypothesized to influence risk

Factors previously evaluated in the literature that were not associated with breast cancer at the p < 0.001 level after adjusting for age, race, and study are shown in Table 5. The factors related to psychological well-being and diet were not found to be associated with breast cancer. Other dietary factors that were

not statistically significant and not included in the table were sugar, carbohydrates, glycemic index, protein, vegetables, dietary fiber, and caffeine. The only variable in the table that was statistically significant at even the p < 0.05 level was the metabolic equivalents of walking times the number of hours of walking per week. This exercise variable included in the table was the most statistically significant of a total of the 23 exercise measures collected by WHI.

Not shown in Table 5 are tests of whether the associations of some factors with breast cancer were influenced by other risk factors. We did not find evidence even at the p < 0.05 level that the association of breast cancer with number of breast aspirations, number of live births, or alcohol consumption were significantly influenced by family history of breast cancer or age. There was also no evidence that the risk of breast cancer associated with estrogen plus progesterone increased as weight increased.

DISCUSSION

A large, diverse study population with long follow-up and

Table 5. Variables of special interest that were not statistically significant at p<0.001 after adjusting for demographic factors

Factor (scale)	All subjects				OS	RCT diet	RCT HT
	% or (mean)	HR†	95% CI	χ²	HR†	HR†	HR†
Psychological							
Religious attendance last month* (1-6)	(2.8)	0.99	0.97-1.01	1.8	0.99	0.99	1.00
Stressful life events* (0-33)	(3.4)	0.97	0.95-1.00	3.5	0.99	0.93	0.93
Optimism* (6-30)	(23.3)	1.00	0.98-1.03	0.1	1.01	0.99	0.98
Depression* (0-1)	(0.04)	1.00	0.97-1.03	0.0	1.01	0.96	1.03
Diet, exercise							
Energy (kcal)*	(1,623)	1.00	0.98-1.03	0.0	1.02	0.98	1.09
Animal protein (g)*	(47.4)	1.00	0.97-1.02	0.1	1.01	0.98	1.21
Folic acid (μg)*	(484.7)	1.00	0.97-1.03	0.0	1.01	0.98	0.83
Fat (g)*	(60.2)	1.00	0.97-1.03	0.0	1.02	0.98	1.18
MET-hours walking a wk*		0.97	0.94-0.99	6.4	0.98	0.94	0.95
0-10	55.0	1.00			1.00	1.00	1.00
10-15	6.5	0.94	0.84-1.04	1.4	0.98	0.93	0.73
> 15	7.2	0.90	0.81-1.00	3.8	0.92	0.80	1.03
Other							
Birth wt*,‡		NA	NA	NA	1.00	NA	NA
FH ovarian cancer	2.3	0.93	0.79-1.09	0.8	0.94	0.98	0.82
Years OC pills*	(2.2)	0.99	0.96-1.01	1.3	0.99	1.00	0.95

HR, hazard ratio; CI, confidence interval; OS, observational study; RCT, randomized controlled trial; RCT HT, combination of two RCTs; Wt, weight; FH, family history; MET, metabolic equivalent; OC, oral contraceptives.

*The hazard ratio was computed for an increase in the variable of one standard deviation; †Hazard ratios were computed for an increase of 1 unit of the ordinal variable or comparison to the reference category of that variable. This reference category has a hazard ratio of 1.00. Hazard ratios were adjusted for all factors independently significant at p < 0.001; ‡Only collected for patients in the observational study.

comprehensive participant information was used to evaluate numerous potential risk factors for breast cancer in postmenopausal women. The study found many factors associated with breast cancer and 20 independently associated at the p < 0.001 level of statistical significance. The large sample size made it possible to detect relatively weak associations and to precisely estimate the strength of association between the risk factor and breast cancer. Because numerous factors were examined, it was possible to show that some factors mediated the risk associated with others. Associations between a risk factor and outcome disappeared after adjusting for mediating risk factors.

The statistically significant independent risk factors in decreasing order of their chi-squared values for one degree of freedom were number of previous breast aspirations, breast cancer in first degree relatives, age, weight, a previous breast biopsy, years using estrogen and progesterone, fewer live births, use of estrogen and progesterone at baseline, age at menopause, a history of thyroid cancer, breast tenderness, digitalis use, at least one drink per week, less restlessness, lack of vaginal dryness, relative with prostate cancer, colon polyps, no breast augmentation, smoking more than 15 cigarettes a day, and no osteoporosis.

Factors of little importance after adjusting for the independent risk factors included socioeconomic status, breast feeding, age at first birth, obesity measures, adult weight gain after adjusting for current weight, and several dietary, psychological, and exercise variables. There was no evidence that family histo-ry influenced the association of breast cancer with the number of live births or alcohol consumption. There was also no evidence that greater weight reduced the risk associated with estrogen plus progesterone.

Greater weight may be associated with breast cancer because it is associated with hormone levels likely to contribute to breast cancer including higher leptin levels, lower adiponectin levels, and higher estrogen levels [2]. Weight is also associated with height, which is a risk factor [3]. The association of a biopsy with breast cancer is thought to be because of the condition that prompted the biopsy [4]. The number of breast aspirations may be an important risk factor because women who had multiple aspirates may have had multiple large cysts, a condition which has been associated with an increased risk of breast cancer for women with a family history of breast cancer [5]. However, we did not find an increase in the association between breast cancer and aspirations for women who had a family history of breast cancer. This increase would be expected if multiple cysts were only a risk factor for women with a family history of breast cancer.

Some variables highly associated with breast cancer in this study directly increase exogenous hormone levels. Others such as number of births and age at menopause indirectly influence hormone levels by influencing the number of lifetime menstrual cycles [6]. The risk associated with digitalis may be because of its estrogenic effect [7], and the protective effect of vaginal dryness and osteoporosis may be because these conditions are

markers for relatively low estrogen levels. Even the protective association of restlessness may be because it indicates lower levels of estrogen although this is not well established [8,9]. Breast augmentation may have a protective association because it is a marker for a smaller breast size, which some studies have suggested is associated with reduced risk [10,11]. In any case there is no evidence from this study that breast implants increase the risk of cancer.

One of the highest hazard ratios was seen with thyroid cancer. This might have been secondary to the radiation exposure in the treatment for thyroid cancer. It is also possible that both the thyroid cancer and the breast cancer may have been increased by medical radiation as treatment for benign conditions such as acne or for medical or dental diagnostic tests that used excessive radiation [12].

Previous epidemiologic studies

A well established set of risk factors commonly used in studies of breast cancer [13] include age at menarche, age at first live birth, number of previous biopsies, and number of first-degree relatives with breast cancer. These risk factors were also found to be statistically significant in the present study although age at menarche was much less significant than others, and age at first birth was not independently significant. Other previously identified risk factors [14] found important in the present study include age, age at menopause, weight, alcohol intake (at least in some studies [15] although not others [16]) hormone therapy, especially estrogen-progestin combinations,and higher bone mineral density [17]. Statistically significant but less important risk factors that had been identified previously include white race [18] and having a male relative with prostate cancer [19]; the latter risk factor has not always been significant when tested [20]. A study using National Cancer Institute's Surveillance, Epidemiology, and End Results data found a much lower risk of breast cancer associated with thyroid cancer, 1.21 [21], than in the present study Another study found that women who had been treated with radioactive iodine had a rate of breast cancer that was 90% greater than in the general population although this increase did not reach statistical significance at the $p < 0.05$ level [22].

Some risk factors found in previous literature were weaker in this study (e.g., exercise [23]) some were not statistically significant (e.g., dietary calories and fat [24], dietary meat [25], and birth weight [26]), and others are entirely explained by the statistically independent risk factors in this study (e.g., adult weight gain [27-30], greater waist girth [28], and socioeconomic status [31,32]). It is possible that the weak association with breast cancer for some risk factors in this study (e.g., age at menarche) was because the variable was inaccurately reported by the participants. We did not find evidence that stressful life events were

risk factors, which agrees with some publications [33] but not others [34]. Our results also did not support previous work showing the influence of a family history of breast cancer on the association between breast cancer and number of children [35] or alcohol [36].

We found evidence of a weak association between breast cancer and smoking that had been found in other studies [37] although in contrast to a previous study [37] we did not find that beginning smoking during the teenage years was an independent risk factor. As with other studies there was no association between breast cancer and diet (including total calories, calories from fat, or folic acid) [38]. Our findings on increased risk associated with breast tenderness and decreased risk associated with hair dyes were opposite of what has been found previously [39,40].

We are unaware of studies that evaluated some risk factors that were independently significant in this study: number of breast aspirations, vaginal dryness, breast augmentation, and restlessness.

This study assessed the relative importance of several risk factors for breast cancer and identified some new risk factors. Some previously identified factors were only weakly associated with breast cancer after adjusting for other risk factors.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare for this study.

REFERENCES

1. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Control Clin Trials 1998;19:61-109.
2. Grossmann ME, Ray A, Nkhata KJ, Malakhov DA, Rogozina OP, Dogan S, et al. Obesity and breast cancer: status of leptin and adiponectin in pathological processes. Cancer Metastasis Rev 2010;29:

641-653.

3. Swanson CA, Coates RJ, Schoenberg JB, Malone KE, Gammon MD, Stanford JL, et al. Body size and breast cancer risk among women under age 45 years. Am J Epidemiol 1996;143:698-706.

4. American Cancer Society. Breast cancer facts & figures 2007-2008 [cited 2013 Jan 18]. Available from: http://www.cancer.org/acs/groups/content/@nho/documents/document/bcfffinalpdf.pdf.

5. Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. N Engl J Med 1985;312:146-151.

6. Clavel-Chapelon F; E3N Group. Cumulative number of menstrual cycles and breast cancer risk: results from the E3N cohort study of French women. Cancer Causes Control 2002;13:831-838.

7. Navab A, Koss LG, LaDue JS. Estrogen-like activity of digitalis: its effect on the squamous epithelium of the female genital tract. JAMA 1965;194:30-32.

8. Hachul H, Baracat EC, Soares JM Jr, Haidar MA, de Mello MT, Tufik S, et al. Estrogen therapy reduces nocturnal periodic limb movements. Maturitas 2007;58:319-322.

9. Polo-Kantola P, Erkkola R, Helenius H, Irjala K, Polo O. When does estrogen replacement therapy improve sleep quality? Am J Obstet Gynecol 1998;178:1002-1009.

10. Egan KM, Newcomb PA, Titus-Ernstoff L, Trentham-Dietz A, Baron JA, Willett WC, et al. The relation of breast size to breast cancer risk in postmenopausal women (United States). Cancer Causes Control 1999;10:115-118.

11. Hsieh CC, Trichopoulos D. Breast size, handedness and breast cancer risk. Eur J Cancer 1991;27(2):131-135.

12. Ron E. Cancer risks from medical radiation. Health Phys 2003;85:47-59.

13. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989;81:1879-1886.

14. McPherson K, Steel CM, Dixon JM. ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. BMJ 2000;321:624-628.

15. Hiatt RA, Klatsky A, Armstrong MA. Alcohol and breast cancer. Prev Med 1988;17:683-685.

16. Harris RE, Spritz N, Wynder EL. Studies of breast cancer and alcohol consumption. Prev Med 1988;17:676-682.

17. Cauley JA, Song J, Dowsett SA, Mershon JL, Cummings SR. Risk factors for breast cancer in older women: the relative contribution of bone mineral density and other established risk factors. Breast Cancer Res Treat 2007;102:181-188.

18. Smigal C, Jemal A, Ward E, Cokkinides V, Smith R, Howe HL, et al. Trends in breast cancer by race and ethnicity: update 2006. CA Cancer J Clin 2006;56:168-183.

19. Beebe-Dimmer JL, Drake EA, Dunn RL, Bock CH, Montie JE, Cooney KA. Association between family history of prostate and breast cancer among African-American men with prostate cancer. Urology 2006;68:1072-1076.

20. Eldon BJ, Jonsson E, Tomasson J, Tryggvadottir L, Tulinius H. Familial risk of prostate cancer in Iceland. BJU Int 2003;92:915-919.

21. Brown AP, Chen J, Hitchcock YJ, Szabo A, Shrieve DC, Tward JD. The risk of second primary malignancies up to three decades after the treatment of differentiated thyroid cancer. J Clin Endocrinol Metab 2008;93:504-515.

22. Goldman MB, Maloof F, Monson RR, Aschengrau A, Cooper DS, Ridgway EC. Radioactive iodine therapy and breast cancer. A follow-up study of hyperthyroid women. Am J Epidemiol 1988;127:969-980.

23. Kruk J. Lifetime physical activity and the risk of breast cancer: a case-control study. Cancer Detect Prev 2007;31(1):18-28.

24. Barrett-Connor E, Friedlander NJ. Dietary fat, calories, and the risk of breast cancer in postmenopausal women: a prospective population-based study. J Am Coll Nutr 1993;12:390-399.

25. Dai Q, Shu XO, Jin F, Gao YT, Ruan ZX, Zheng W. Consumption of animal foods, cooking methods, and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2002;11:801-808.

26. Xu X, Dailey AB, Peoples-Sheps M, Talbott EO, Li N, Roth J. Birth weight as a risk factor for breast cancer: a meta-analysis of 18 epidemiological studies. J Womens Health (Larchmt) 2009;18:1169-1178.

27. Ahn J, Schatzkin A, Lacey JV Jr, Albanes D, Ballard-Barbash R, Adams KF, et al. Adiposity, adult weight change, and postmenopausal breast cancer risk. Arch Intern Med 2007;167:2091-2102.

28. Krebs EE, Taylor BC, Cauley JA, Stone KL, Bowman PJ, Ensrud KE. Measures of adiposity and risk of breast cancer in older postmenopausal women. J Am Geriatr Soc 2006;54:63-69.

29. Eliassen AH, Colditz GA, Rosner B, Willett WC, Hankinson SE. Adult weight change and risk of postmenopausal breast cancer. JAMA 2006;296:193-201.

30. Lahmann PH, Lissner L, Gullberg B, Olsson H, Berglund G. A prospective study of adiposity and postmenopausal breast cancer risk: the Malmö Diet and Cancer Study. Int J Cancer 2003;103:246-252.

31. Heck KE, Wagener DK, Schatzkin A, Devesa SS, Breen N. Socioeconomic status and breast cancer mortality, 1989 through 1993: an analysis of education data from death certificates. Am J Public Health 1997;87:1218-1222.

32. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes Control 2001;12:703-711.

33. Achat H, Kawachi I, Byrne C, Hankinson S, Colditz G. A prospective study of job strain and risk of breast cancer. Int J Epidemiol 2000;29:622-628.

34. Eskelinen M, Ollonen P. Life stress due to losses and deficit in childhood and adolescence as breast cancer risk factor: a prospective case-control study in Kuopio, Finland. Anticancer Res 2010;30:4303-4308.

35. Colditz GA, Rosner BA, Speizer FE. Risk factors for breast cancer according to family history of breast cancer. For the Nurses' Health Study Research Group. J Natl Cancer Inst 1996;88:365-371.

36. Vachon CM, Cerhan JR, Vierkant RA, Sellers TA. Investigation of an interaction of alcohol intake and family history on breast cancer risk in the Minnesota Breast Cancer Family Study. Cancer 2001;92:240-248.

37. Luo J, Margolis KL, Wactawski-Wende J, Horn K, Messina C, Stefanick ML, et al. Association of active and passive smoking with risk of breast cancer among postmenopausal women: a prospective cohort study. BMJ 2011;342:d1016.

38. Michels KB, Mohllajee AP, Roset-Bahmanyar E, Beehler GP, Moysich KB. Diet and breast cancer: a review of the prospective observational studies. Cancer 2007;109:2712-2749.

39. Khan SA, Apkarian AV. Mastalgia and breast cancer: a protective association? Cancer Detect Prev 2002;26:192-196.

40. Shore RE, Pasternack BS, Thiessen EU, Sadow M, Forbes R, Albert RE. A case-control study of hair dye use and breast cancer. J Natl Cancer Inst 1979;62:277-283.