suno Y, Shimosato Y, Hirohashi S. Application of the p53 gene mutation pattern for differential diagnosis of primary versus metastatic lung carcinomas. Diagn Mol Pathol 1993;2: 29–35.

*(38)* Johnson BE, Linnoila RI, Williams JP, Venzon DJ, Okunieff P, Anderson GB, et al. Risk of second aerodigestive cancers increases in patients who survive free of small-cell lung cancer for more than 2 years. J Clin Oncol 1995; 13:101–11.

*(39)* Fisher JC. Multiple mutation theory of carcinogenesis. Nature 1958;181:651–2.

*(40)* Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. Cancer Res 1991;51:3075–9.

*(41)* Loeb LA. Microsatellite instability: marker of a mutator phenotype in cancer. Cancer Res 1994;54:5059–63.

*(42)* O'Connell P, Pekkel V, Fuqua S, Osborne CK, Allred DC. Molecular genetic studies of early breast cancer evolution. Breast Cancer Res Treat 1994;32:5–12.

*(43)* Wright TC, Kurman RJ, Ferenczy A. Precancerous lesions of the cervix. In: Kurman RJ, editor. Blaustein's pathology of the female genital tract. New York: Springer-Verlag, 1994:229–77.

*(44)* Fearon ER. Molecular genetic studies of the adenoma–carcinoma sequence. Adv Intern Med 1994;39:123–47.

*(45)* D'Amico D, Carbone DP, Johnson BE, Meltzer SJ, Minna JD. Polymorphic sites within the MCC and APC loci reveal very frequent loss of heterozygosity in human small cell lung cancer. Cancer Res 1992;52:1996–9.

*(46)* Wieland I, Bohm M. Frequent allelic deletion at a novel locus on chromosome 5 in human lung cancer. Cancer Res 1994;54:1772–4.

## Notes

Supported by Public Health Service contract N01CN45580-01 and grant 1P50CA70907–01 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services.

We thank Jing Xu for assistance with the statistical analysis.

Manuscript received April 21, 1997; revised July 15, 1997; accepted July 16, 1997.

# Geographic Variation in Breast Cancer Incidence Rates in a Cohort of U.S. Women

*Francine Laden, Donna Spiegelman, Lucas M. Neas, Graham A. Colditz, Susan E. Hankinson, JoAnn E. Manson, Celia Byrne, Bernard A. Rosner, Frank E. Speizer, David J. Hunter**

***Background:*** **Breast cancer mortality and incidence rates vary by geographic region in the United States. Previous analytic studies have measured mortality, not incidence, and have used regional prevalences to control for geographic variation in risk factors rather than adjusting for risk factors measured at the level of the individual. We prospectively evaluated regional variation in breast cancer incidence rates in the Nurses' Health Study and assessed the influence of breast cancer risk factors measured at the individual level.** ***Methods:*** **The Nurses' Health Study cohort was established in 1976 when 121 700 female nurses aged 30–55 years living in 11 U.S. states were enrolled. These states represent all four regions of the continental United States. We identified 3603 incident cases of invasive breast cancer through 1992 (1 794 565 person-years of follow-up). We calculated relative risks (RRs) adjusted for age and for age and established risk factors (i.e., multivariate-adjusted analysis), comparing California, the Northeast, and the Midwest with the South.** ***Results:*** **For premenopausal women, there was little evidence of regional variation in breast cancer incidence rates, either in age-adjusted or in multivariate-adjusted analyses. For postmenopausal women in California, age-adjusted risk was modestly elevated (RR = 1.24; 95% confidence interval [CI] = 1.05–1.47); after adjusting for age and for established risk factors, the excess rate in California was attenuated by 25% (RR = 1.18; 95% CI = 1.00–1.40). No excess of breast cancer incidence was observed for postmenopausal women in either the Northeast or the Midwest.** ***Conclusions:*** **Little regional variation in age-adjusted breast cancer incidence rates was observed, with the exception of a modest excess for postmenopausal women in California. Adjustment for differences in the distribution of established risk factors explained some of the excess risk in California. [J Natl Cancer Inst 1997; 89:1373–8]**

Mortality rates from breast cancer vary by geographic region within the United States. Average annual age-adjusted mortality rates ranged from 18.0 per 100 000 women in Hawaii to 35.7 per 100 000 women in the District of Columbia during the period 1987 through 1991 *(1)*. In regional analyses, mortality rates among women older than 50 in the northeastern region have been reported to be from 20% to 50% higher than those in the southern United States *(2–4)*. This pattern, along with the rise in incidence rates

**Affiliations of authors:* F. Laden, S. E. Hankinson, Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, and Department of Epidemiology, Harvard School of Public Health, Boston; D. Spiegelman, Departments of Epidemiology and Biostatistics, Harvard School of Public Health; L. M. Neas, Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, and Departments of Epidemiology and Environmental Health, Harvard School of Public Health; G. A. Colditz, D. J. Hunter, Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, and Department of Epidemiology, Harvard School of Public Health, and Harvard Center for Cancer Prevention; J. E. Manson, Channing Laboratory, Department of Medicine, and Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School; C. Byrne, Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School; B. A. Rosner, Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, and Department of Biostatistics, Harvard School of Public Health, Boston; F. E. Speizer, Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, and Department of Environmental Health, Harvard School of Public Health.

*Correspondence to:* Francine Laden, M.S., Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, 181 Longwood Ave., Boston, MA 02115.

*See* ''Notes'' following ''References.''

throughout most of this century *(5)*, evidence that higher rates are found in urban areas *(2,6)*, and reports of suspected cancer clusters *(7–9)* have led to speculation about environmental pollutants as causes of breast cancer. However, there is also evidence that established risk factors for breast cancer, such as fertility rates *(10)*, delayed childbearing *(2,4)*, economic status and educational level *(11)*, exogenous hormone use *(4)*, and alcohol consumption *(12)*, vary modestly between regions. Furthermore, regional differences in mortality could be due to differences in the prevalence of early detection [e.g., mammographic screening practices vary modestly by region *(13)*] and/or treatment of incident breast cancers *(14–16)*. In most previous nationwide studies, mortality rates have been used, and geographic variation of potentially confounding factors was controlled for by using regional prevalences of these factors, rather than adjusting for individually measured risk factors.

We evaluated prospectively the regional variation of invasive breast cancer incidence in the Nurses' Health Study controlling for breast cancer risk factors collected at the individual level. The Nurses' Health Study represents a single occupational group; potential confounders related to socioeconomic status that are notoriously difficult to adjust for directly are at least partially removed by restricting to this narrower socioeconomic stratum.

## Methods

### Study Population

The Nurses' Health Study is an ongoing prospective cohort study established in 1976 when 121 700 registered nurses completed a mailed questionnaire that included items about risk factors for breast cancer and other diseases. At enrollment, the participants were between the ages of 30 and 55 years old and resided in 11 large states (California, Connecticut, Florida, Maryland, Massachusetts, Michigan, New Jersey, New York, Ohio, Pennsylvania, and Texas). These states were originally chosen based on their size and the approval of the study by the state nursing associations. No restrictions were made on the basis of ethnicity or race; however, the participants were primarily Caucasian (approximately 97%), reflecting the ethnic background of women trained as registered nurses. Every 2 years, participants completed follow-up questionnaires to update information on risk factors for breast cancer and to report the occurrence of breast cancer and other illnesses.

The subjects included in this analysis were the 118 349 women who did not report breast cancer or other cancers (with the exception of nonmelanoma skin cancer) at baseline in 1976.

### Assessment of Exposure

Place of residence was defined as the region of the country in which the participant lived in 1976, the beginning of the study. We grouped the 11 states represented by this cohort into four regions based on census definitions *(17)*—Northeast (Connecticut, Massachusetts, New York, New Jersey, and Pennsylvania), Midwest (Michigan and Ohio), South (Florida, Maryland, and Texas), and West (California). In 1976, 14 674 (12.4%) of the cohort lived in California, 68 921 (58.2%) lived in the Northeast, 21 702 (18.3%) lived in the Midwest, and 13 052 (11.0%) lived in the South. The nurses lived throughout the 11 states; the county-specific population distribution of the cohort reflected that of the general population of white women of the same age range, with the exception of some underascertainment of women in large urban counties in the Northeast and in small counties in the South. To account for duration of exposure, we also defined residence by region in which the participant lived in 1986, 10 years after the start of the study. In some analyses, we restricted the cohort to the 110 741 participants (94%) who lived in the same region in 1976 and in 1986, thereby defining a stable population. In further analyses, we restricted to the 62 672 women (53%) who reported in 1992 that they had lived in the same region at birth, at age 15 years, and at age 30 years.

### Assessment of Outcome

Diagnoses of breast cancer were reported on the biennial follow-up questionnaires. We attempted to contact nonrespondents by telephone and identified deaths through next of kin or searches of the National Death Index. For each case of breast cancer reported, we requested permission to obtain medical records and pathology reports to confirm the diagnosis. Because the accuracy of self-reported breast cancer was extremely high *(18)*, we included in this report the small number of cases for whom pathology reports were not obtained (n = 191). In the majority of analyses, we considered incident cases of invasive breast cancer only. In one analysis, we included incident cases of *in situ* carcinoma of the breast, and we also analyzed breast cancer mortality, as determined by review of death certificates and medical records.

### Assessment of Breast Cancer Risk Factors

We obtained information on known and suspected risk factors for breast cancer in 1976 and updated the information at the beginning of each 2-year period, as appropriate. We included the following risk factors in the multivariate models: age, menopausal status, age at menopause, age at menarche, parity, age at first full-term pregnancy, use of oral contraceptives, use and duration of use of postmenopausal hormone therapy, history of breast cancer in a mother or a sister, history of benign breast disease, height, current body mass index (weight [kg]/height$[m]^2$), and body mass index at age 18 years. We classified a woman as postmenopausal from the time she returned a questionnaire on which she reported natural menopause or hysterectomy with bilateral oophorectomy. Women reporting hysterectomy without bilateral oophorectomy were assumed to be postmenopausal at the age when natural menopause had occurred in 90% of the cohort (54 years for current cigarette smokers and 56 years for nonsmokers); otherwise, we considered them to be of uncertain menopausal status.

We collected information on alcohol consumption prospectively beginning in 1980 *(19)*, and it was updated in 1984, 1986, and 1990. In 1988 we inquired as to whether the participant had ever had a mammographic examination.

### Allocation of Person-Time

Follow-up began on June 1, 1976. Each participant contributed person-time to the analysis up until June 1, 1992, until date of diagnosis of breast cancer (date of death from breast cancer for the mortality analysis), or until the date of death from other causes, whichever came earlier. In all analyses, except those involving mortality, women who reported a diagnosis of cancer other than nonmelanoma skin cancer on any questionnaire were excluded from subsequent follow-up at the beginning of the next follow-up cycle. For analyses adjusting for alcohol intake, we began follow-up in 1980 (when alcohol consumption was first assessed) and limited the cohort to the women who were cancer-free at the start of the 1980 follow-up and who provided detailed dietary information in 1980 (n = 89 512). Person-time for each participant was allocated to their region of residence in 1976. Each individual's risk factor status was updated at the beginning of each 2-year period on the basis of information provided on the follow-up questionnaires. The follow-up rate was similar between regions and averaged 95% of potential person-time.

We performed all analyses within the entire cohort and separately among premenopausal and postmenopausal women. In 1976, 22 990 women reported that they were postmenopausal and entered the postmenopausal follow-up in the period 1976 through 1978. As women became postmenopausal during follow-up, their person-time was added to the postmenopausal analysis. By the start of the 1990 through 1992 time period, 71 070 women were defined as postmenopausal. Women who started follow-up as premenopausal (84 692 in 1976) were excluded from the premenopausal analysis as their menopausal status changed. Women with missing or uncertain menopausal status during a given time period were excluded from the stratified analysis during that time period.

### Regional Distribution of Risk Factors

To assess the regional distribution of breast cancer risk factors and their potential to confound the region/breast cancer relationship, we calculated the proportion of person-time in each covariate category by menopausal status, standardized to the age distribution of the premenopausal or postmenopausal cohort. For the risk factors assessed for the full period of follow-up, we used the age distribution of the entire postmenopausal cohort and of the entire premenopausal cohort to standardize the postmenopausal and premenopausal prevalences, respectively. For alcohol use, we used the age distributions of the 1980 cohort with dietary data. History of mammography was first asked in 1988. Therefore, we calcu-

lated the percent of women who answered the 1988 questionnaire and reported ever having had a mammographic examination.

### Incidence Rates and Comparison to National Rates

To calculate the age-standardized incidence rates, we divided the number of incident breast cancers by the person-time of follow-up and standardized the regional incidence rates to the age distribution of the entire cohort at baseline. To assess the comparability of the Nurses' Health Study breast cancer incidence with the national incidence rates of invasive breast cancer, we calculated the expected number of cases using age-specific incidence rates observed by the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER)[1] Program over the period 1976 through 1990 *(20)*, standardized to the age distribution of the Nurses' Health Study. The SEER program consists of data from nine population registries for cancer incidence in various locations that represent approximately 10% of the U.S. population.

### Multivariate Analyses

We used the likelihood ratio test, comparing the model with indicator variables for both age and region with the model with only age, to evaluate the contribution of region to the model and address the general question of whether regional variation existed in this cohort. To be consistent with previously published research from the NCI, we chose the women residing in the South as the reference group when comparing incidence between regions *(4)*. We calculated relative risks (RRs), dividing the incidence rate in each region by the incidence rate in the South. To control simultaneously for potential confounding factors we conducted proportional hazard analyses *(21)* by using a pooled logistic regression model *(22,23)* with indicator variables for each region for each category of each breast cancer risk factor and for 2-year intervals of calendar time. We calculated the 95% confidence intervals (CIs) for each RR.

To test whether including the small numbers of non-Caucasians in the analysis altered our conclusions, we compared our overall results with results obtained by restricting the cohort to white women who answered a question on ethnicity (n = 95 672) in 1992.

## Results

### Distribution of Risk Factors in the Cohort

The age-standardized prevalences of established and potential risk factors for breast cancer varied modestly across region for both premenopausal and postmenopausal women. (These differences were statistically significant; however, even very small differences are statistically significant with such a large sample size.) Prevalences for postmenopausal women are shown in Table 1; prevalences for premenopausal women were similar. Women residing in California were more likely to delay childbearing compared with women in other regions, and women in California and the South had slightly fewer children than women in the Northeast and Midwest. Women in California were much more likely to use oral contraceptives and postmenopausal hormones, to have had a mammographic examination, and to consume alcohol than other women. Women in the Midwest and Northeast, in general, were heavier than other women both at age 18 years and currently. Age at menopause varied by region; southern women were slightly younger at menopause than other women. In summary, the group residing in California had the highest prevalence of established breast cancer risk factors and women in the Midwest had the lowest.

**Table 1.** Age-standardized distribution* of breast cancer risk factors for postmenopausal women by region

| | California, % | Northeast, % | Midwest, % | South, % |
|---|---|---|---|---|
| Menarche, ≤12 y | 45.8 | 47.0 | 47.0 | 45.8 |
| Nulliparous | 10.1 | 7.9 | 7.4 | 9.5 |
| Parity, ≥5 (among parous women) | 15.8 | 21.2 | 22.3 | 13.9 |
| ≤24 y at first birth (among parous women) | 43.2 | 49.8 | 53.7 | 51.7 |
| ≥30 y at first birth (among parous women) | 15.5 | 12.0 | 10.0 | 10.3 |
| Ever use of oral contraceptives | 37.7 | 29.3 | 36.3 | 31.6 |
| Current use of postmenopausal hormones, ≥5 y | 19.9 | 7.9 | 13.1 | 18.1 |
| History of benign breast disease | 29.6 | 26.5 | 28.8 | 31.1 |
| Family history in mother or sister | 8.8 | 8.6 | 8.2 | 8.3 |
| Height, ≥168 cm | 23.8 | 19.4 | 20.3 | 23.6 |
| Body mass index at age 18 y, <19.0 | 14.5 | 11.5 | 12.5 | 13.8 |
| Body mass index at age 18 y, ≥24.0 | 12.8 | 17.0 | 17.8 | 14.5 |
| Current body mass index, <21.0 | 14.6 | 11.2 | 11.7 | 13.4 |
| Current body mass index, ≥29.0 | 12.1 | 15.2 | 16.5 | 12.8 |
| Alcohol, ≥15 g/day | 14.3 | 10.7 | 8.0 | 9.9 |
| Ever mammogram by 1988 | 83.8 | 74.8 | 77.6 | 77.5 |
| Age at menopause, <40 y | 13.4 | 11.5 | 11.6 | 15.9 |
| Age at menopause, ≥50 y | 41.3 | 41.5 | 40.7 | 34.9 |

*Percents represent the age-adjusted person-time allocated to that category divided by the total person-time of follow-up for the region.

### Invasive Breast Cancer

Between 1976 and 1992, 3603 incident cases of invasive breast cancer occurred among 118 349 nurses during 1 794 565 person-years of follow-up. The overall age-adjusted incidence rate was 200.8 cases per 100 000 person-years. As defined by residence at baseline in 1976, the region-specific incidence rates (per 100 000 person-years), standardized to the age distribution of the entire cohort, were 225.1 in California, 197.9 in the Northeast, 196.5 in the Midwest, and 193.4 in the South. The overall incidence rate observed in the Nurses' Health Study was 8% higher than the expected incidence rate calculated using the SEER program incidence rates for white women during a similar period (1976 through 1990).

Among all women, region contributed significantly to the age-adjusted model ($P = .05$). We observed a small, but statistically significant, elevation of the age-adjusted breast cancer incidence in California compared with the South (RR = 1.16; 95% CI = 1.02–1.32) (Table 2). However, the incidence rates in the Northeast and the Midwest were not elevated relative to the South. After adjusting for established breast cancer risk factors, the contribution of region to the model was no longer significant. However, the RR for California was only slightly attenuated and still of borderline significance (RR = 1.13; 95% CI = 0.99–1.29). The RRs for the Northeast and the Midwest were similar to the age-adjusted values. Including cases of *in situ* carcinoma of the breast along with invasive breast cancer cases did not notably change the age-adjusted RRs. The RRs for the established breast cancer risk factors were consistent with results from previous reports *(24)*.

During the period of follow-up, 1196 premenopausal women and 2005 postmenopausal women developed invasive breast cancer. For the premenopausal women, there was little evidence of regional variation in either the age-adjusted or multivariate-adjusted analyses (Table 2). For the postmenopausal women, we observed a statistically significant el-

**Table 2.** Relative risk (RR) of invasive breast cancer incidence in relation to region of residence in the United States, by menopausal status, among 118 349 women aged 30–55 years in 1976 and followed through 1992

| | California | Northeast | Midwest | South |
|---|---|---|---|---|
| | | ***All women**** | | |
| No. of cases | 535 | 2034 | 639 | 395 |
| Person-years of observation | 220 476 | 1 048 085 | 329 008 | 196 996 |
| RR (age adjusted) (95% CI) | 1.16 (1.02–1.32) | 1.02 (0.92–1.14) | 1.02 (0.90–1.15) | 1.00 |
| RR (multivariate) (95% CI)† | 1.13 (0.99–1.29) | 1.05 (0.94–1.17) | 1.03 (0.91–1.17) | 1.00 |
| | | ***Premenopausal women*** | | |
| No. of cases | 142 | 717 | 223 | 114 |
| Person-years of observation | 83 093 | 489 300 | 152 206 | 75 740 |
| RR (age adjusted) (95% CI) | 1.07 (0.83–1.36) | 0.98 (0.80–1.19) | 0.99 (0.79–1.23) | 1.00 |
| RR (multivariate) (95% CI)‡ | 1.02 (0.80–1.31) | 1.01 (0.83–1.23) | 1.02 (0.81–1.28) | 1.00 |
| | | ***Postmenopausal women*** | | |
| No. of cases | 327 | 1103 | 353 | 222 |
| Person-years of observation | 103 956 | 420 610 | 134 418 | 90 585 |
| RR (age adjusted) (95% CI) | 1.24 (1.05–1.47) | 1.08 (0.93–1.24) | 1.08 (0.91–1.27) | 1.00 |
| RR (multivariate) (95% CI)§ | 1.18 (1.00–1.40) | 1.12 (0.97–1.30) | 1.09 (0.92–1.29) | 1.00 |

*Women of uncertain menopausal status were included in analyses of all women but were excluded from the stratified analyses.

†Multivariate RR and 95% confidence interval (CI), adjusted for age in 5-year categories, age at menarche (⩽12, 13, or ⩾14 years), parity (nulliparous, 1–2, 3–4, or ⩾5), age at first birth (nulliparous, ⩽24, 25–29, or ⩾30 years), use of oral contraceptives (ever or never), menopausal status (premenopausal, postmenopausal, or unknown), use and duration of use of postmenopausal hormones (never use, current use <5 years, current use ⩾5 years, or past use), history of breast cancer in a mother or sister, history of benign breast disease, and body mass index (five groups).

‡Multivariate RR and 95% CIs, adjusted for same risk factors as in full cohort analysis, except menopausal status and postmenopausal hormone use, are excluded from models.

§Multivariate RR and 95% CIs, adjusted for same risk factors as full cohort analysis, with the addition of age at menopause in 2-year categories.

evated age-adjusted incidence rate in California (RR = 1.24; 95% CI = 1.05–1.47). The age-adjusted RR in both the Northeast and Midwest was 1.08, and neither was statistically significant. After adjusting for all of the breast cancer risk factors, the excess rate in California was attenuated by 25% (RR = 1.18) but remained of borderline significance (95% CI = 1.00–1.40). The strongest confounding factors were age at first birth, postmenopausal hormone use, and age at menopause. Finer categories of duration of postmenopausal hormone use did not change the association between region and breast cancer risk. Controlling for type of menopause (natural, surgical, or other) also did not alter the association. The increased risk in California was apparent in both the northern and southern halves of the state.

In age-adjusted mortality analyses based on 82 deaths in California and 52 deaths in the South that occurred among postmenopausal women who were cancer free at baseline, risk of breast cancer death was nonsignificantly higher in California, RR = 1.34 (95% CI = 0.95–1.89).

Ninety-four percent of the cohort lived in the same region in 1986 as they did in 1976. When we restricted the cohort to these women, the RRs for invasive breast cancer incidence were comparable to those obtained with the full cohort. Additionally, we restricted the cohort to women who lived in the same region at birth, at age 15 years, at age 30 years, and in 1976. The results were similar to those obtained using the full cohort, except for a stronger association observed for California (Table 3).

Adjusting for alcohol intake for the

**Table 3.** Age-adjusted relative risk (RR) of invasive breast cancer incidence (95% confidence interval [CI]) for region, restricted to women who lived in the same region throughout their lifetime:* follow-up 1976–1992

| | California | Northeast | Midwest | South |
|---|---|---|---|---|
| | | ***All women†*** | | |
| No. of cases | 92 | 1123 | 335 | 113 |
| Person-years of observation | 41 248 | 658 713 | 197 572 | 70 579 |
| RR (age-adjusted) (95% CI) | 1.37 (1.04–1.80) | 1.10 (0.91–1.34) | 1.09 (0.88–1.35) | 1.00 |
| RR (multivariate) (95% CI)‡ | 1.35 (1.02–1.78) | 1.14 (0.94–1.39) | 1.13 (0.91–1.40) | 1.00 |
| | | ***Premenopausal women*** | | |
| No. of cases | 31 | 408 | 119 | 36 |
| Person-years of observation | 18 705 | 331 033 | 98 348 | 29 894 |
| RR (age-adjusted) (95% CI) | 1.33 (0.82–2.14) | 1.00 (0.71–1.40) | 1.00 (0.69–1.46) | 1.00 |
| RR (multivariate) (95% CI)‡ | 1.27 (0.79–2.05) | 1.04 (0.74–1.47) | 1.05 (0.72–1.52) | 1.00 |
| | | ***Postmenopausal women*** | | |
| No. of cases | 50 | 617 | 186 | 63 |
| Person-years of observation | 17 938 | 268 672 | 80 436 | 31 966 |
| RR (age-adjusted) (95% CI) | 1.34 (0.92–1.94) | 1.16 (0.89–1.50) | 1.17 (0.88–1.55) | 1.00 |
| RR (multivariate) (95% CI)‡ | 1.31 (0.90–1.91) | 1.22 (0.94–1.59) | 1.19 (0.89–1.59) | 1.00 |

*As defined by living in the same region at birth, age 15 years, and age 30 years. This question was asked in 1992, therefore only women who answered that questionnaire were eligible for this analysis.

†Women of uncertain menopausal status were included in analyses of all women but were excluded from the stratified analyses.

‡Multivariate RR and 95% CIs, adjusted for breast cancer risk factors as described in Table 2.

time period 1980 through 1992 only slightly attenuated the RR comparing California to the South; multivariate RR = 1.19 (95% CI = 0.97–1.46) with alcohol versus RR = 1.21 (95% CI = 0.99–1.48) without alcohol. Adjustment for alcohol did not alter the lack of association between the other regions and the risk of breast cancer. California had the largest percentage of non-Caucasian population (8%). However, restricting the cohort to white women defined as those who did not report Hispanic, African-American, or Asian ancestry did not materially change the RRs. Results were also similar when we restricted the cohort to women who had had at least one mammographic examination.

## Discussion

In prospective analyses of a socioeconomically restricted cohort with members drawn from all four U.S. census-defined regions, we did not observe the hypothesized elevated rate of breast cancer incidence in the Northeast compared with the South nor did we see a significant elevation in the Midwest. Premenopausal breast cancer incidence did not vary significantly by region. We observed a marginally statistically significant elevated age-adjusted breast cancer incidence rate in California among postmenopausal women that increased slightly when we restricted the analysis to women who had lived in the same region throughout most of their lives. In this cohort, the South had a slightly higher prevalence of risk factors for breast cancer compared with the Northeast and Midwest for both premenopausal and postmenopausal women and a slightly lower prevalence of risk factors compared with California. After controlling for these factors, 25% of the excess rate of postmenopausal breast cancer in California was explained.

Our results were consistent in direction, although not in magnitude, with previous mortality studies. Sturgeon et al. *(4)* used data from the National Center for Health Statistics from 1987 and observed elevated age-adjusted mortality rate ratios in all regions compared with the South for women aged 50–79 years. The RRs were 1.15 in the West, 1.30 in the Northeast, and 1.18 in the Midwest. In an ecological analysis controlling for group-defined risk and prognostic factors, they were able to explain 50% of the excess mortality in the Northeast and Midwest, but only 10% in the West *(4)*. Blot et al. *(2)* observed a 20% increased rate of breast cancer death in large counties of the Northeast compared with large counties of the South and a 50% increase when they compared small counties. The excess risk in the West ranged from 7% to 30% in large and small counties, respectively. The use of 1960 census data to control for income, urbanization, birth rate, and German or Scandinavian ancestry did not eliminate the region effect among older women *(2)*.

A limitation of our study is that we used only 11 states to make inferences about four large regions. However, these states contain 53% of the entire U.S. population and account for large proportions of the populations of their respective regions *(17)*. With the exception of an under-ascertainment of women in large urban counties in the Northeast and in small counties in the South, the county-specific geographic distribution of participants in the Nurses' Health Study is remarkably representative of these states (Laden F, Neas LM, Hunter DJ: unpublished data). However, these states are not necessarily representative of the region as a whole, particularly in the South and West. For example, the proportion of college-educated persons in Florida, Maryland, and Texas is closer to the proportion observed in the northern states than to the proportion in the remainder of the southern region *(11)*. This limitation may explain why we did not observe RRs for the Northeast and Midwest of the same magnitude as seen in the previous mortality studies. Furthermore, using only California to represent the West may explain why our RR in the West is higher than previously observed. The breast cancer mortality rate in California (1986–1990) was the highest in the Western region and the San Francisco SEER registry reports the highest incidence rate of all registries in the nation *(1)*.

The fact that in most states the geographic distribution of nurses by county was similar to that of white women in general suggests that we did not fail to detect an elevation of risk in certain states due to underrepresentation of individual counties in which breast cancer rates may be higher. We cannot exclude the possibility, however, that very localized exposures within counties might cause breast cancer and be more common in some states than others or that within each county nurses were systematically less likely to live near these sources of exposure. Our results do diminish the likelihood that environmental exposures that are widespread and differ between regions cause large differences in breast cancer rates.

Another potential limitation of this study is that we did not have prospective information on screening. However, the prevalence of mammography was high and similar in all regions, suggesting that differential mammography rates were unlikely to have had substantial influence on the results. Results were similar when we restricted the analysis to women who had had at least one mammographic examination.

Residual confounding could be responsible for our inability to explain some of the excess age-adjusted rate of breast cancer in California compared with the South. We were not able to control for potential risk factors such as physical activity, diet, or alcohol consumption in early life. We did not directly measure hypothesized environmental risk factors for breast cancer, such as reduced sunlight *(25)*, electromagnetic fields *(26,27)*, exposure to organochlorine compounds *(28,29)*, and other pollutants *(30)*. Thus, our results do not rule out the possibility that differences in exposure to these factors between California and the rest of the country might be responsible for some of the small residual difference in breast cancer incidence that was observed. Reassuringly, a recent study showed that regional differences in known breast cancer risk factors completely accounted for the modest elevation in breast cancer incidence rates in the San Francisco Bay Area compared with seven other SEER registries *(31)*.

Use of the Nurses' Health Study cohort restricts the study population to one occupational group of mostly Caucasian women. Thus, the range of possible occupational exposures is reduced, limiting the generalizability of the study. However, this restriction allows us to focus on nonoccupational environmental exposures that might be associated with region. Also, because the Nurses' Health Study is relatively homogeneous com-

pared with the general population, we indirectly controlled for potential confounding by socioeconomic status, and the participants' relatively good access to health care should reduce potential confounding by regional differences in early diagnosis. Aspects of socioeconomic status that vary greatly by region in the general population may explain why we did not see the same magnitude of regional variation reported in previous studies.

Despite these limitations, this study assesses nationwide variation of breast cancer incidence rates in a prospective analysis using risk factors assessed at the individual instead of the group level. The use of incidence, as opposed to mortality rates, avoids bias from potential regional differences in early detection and treatment effectiveness as well as possible differential migration among cases of breast cancer due to health care concerns or retirement. Our results suggest that there is a small excess age-adjusted incidence of postmenopausal breast cancer in California but not in the Northeast or Midwest. Some of the excess rate in California can be explained by established risk factors. Geographic variation in breast cancer rates at the state or regional level is unlikely to be due to region-specific differences in exposures to widespread nonoccupational environmental pollutants.

## References


*(1)* Ries LA, Miller BA, Hankey BF, Kosary CL, Harras A, Edwards BK, editors. SEER cancer statistics review, 1973–1991: tables and graphs, Bethesda (MD): National Cancer Institute; 1994; NIH Publ No. 94–2789.

*(2)* Blot WJ, Fraumeni JF Jr. Stone BJ. Geographic patterns of breast cancer in the United States. J Natl Cancer Inst 1977;59:1407–11.

*(3)* Pickle LW, Mason TJ, Howard N, Hoover R, Fraumeni JF Jr. Atlas of U.S. cancer mortality among whites: 1950–1980. Bethesda (MD): National Institutes of Health, 1987; DHHS Publ No. (NIH)87–2900.

*(4)* Sturgeon SR, Schairer C, Gail M, McAdams M, Brinton LA, Hoover RN. Geographic variation in mortality from breast cancer among white women in the United States. J Natl Cancer Inst 1995;87:1846–53.

*(5)* Kelsey JL, Horn-Ross PL. Breast cancer: magnitude of the problem and descriptive epidemiology. Epidemiol Rev 1993;15: 7–16.

*(6)* Nasca PC, Mahoney MC, Wolfgang PE. Population density and cancer incidence differentials in New York State, 1978–82. Cancer Causes Control 992;3:7–15.

*(7)* New York State Dept of Health, Dept of Community and Preventive Medicine SUNY at Stonybrook, Nassau County Dept of Health, Suffolk County Dept of Health Services. The Long Island Breast Cancer Study: Report Number 1, 1988.

*(8)* Brody JG, Rudel R, Maxwell NI, Swedes SR. Mapping out a search for environmental causes of breast cancer. Public Health Reports 1996; 11:494–507.

*(9)* Dusich K, Sigurdson E, Hall WN, Dean AG. Minnesota Dept. of Health: Cancer rates in a community exposed to low levels of creosote components in municipal water. Min Med 1980;63:803–6.

*(10)* Clarke SC, Ventura SJ. Birth and fertility rates for states: United States, 1990. National Center for Health Statistics. Vital Health Stat 21(52), 1994.

*(11)* USA Counties on CD-ROM— prepared by the Bureau of the Census.—Washington (DC): The Bureau [producer and distributor], 1994.

*(12)* Piani A, Schoenborn C. Health promotion and disease prevention United States, 1990. National Center for Health Statistics. Vital Health Stat 10(185), 1993.

*(13)* Dawson DA, Thompson GB. Breast cancer risk factors and screening: United States, 1987. National Center for Health Statistics. Vital Health Stat 10*(172).* 1990.

*(14)* Ayanian JZ, Guadagnoli E. Variations in breast cancer treatment by patient and provider characteristics. Breast Cancer Res Treat 1996;40: 65–74.

*(15)* Choi WS, Parker BA, Pierce JP, Greenberg ER. Regional differences in the incidence and treatment of carcinoma *in situ* of the breast. Cancer Epidemiol Biomarkers Prev 1996;5: 317–20.

*(16)* Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma *in situ* of the breast. JAMA 1996;275:913–8.

*(17)* U.S. Department of Commerce, Bureau of the Census. A Guide to State and Local Census Geography. Princeton (NJ): Association of Public Data Users, 1993.

*(18)* Colditz GA, Martin P, Stampfer MJ, Willett WC, Sampson L, Rosner B, et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. Am J Epidemiol 1986;123: 894–900.

*(19)* Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122: 51–65.

*(20)* SEER 1973–90. Cancer Incidence Public Use Database.

*(21)* Cox DR. Regression models and life-tables. J R Stat Soc (B) 1972;34:187–220.

*(22)* Abbott RD. Logistic regression in survival analysis. Am J Epidemiol 1985;121:465–71.

*(23)* D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. Stat Med 1990;9:1501–15.

*(24)* Colditz GA. Epidemiology of breast cancer: findings from the Nurses' Health Study. Cancer 1993;71(4 Suppl):1480–89.

*(25)* Garland FC, Garland CF, Gorham ED, Young JF. Geographic variation in breast cancer mortality in the United States: a hypothesis involving exposure to solar radiation. Prev Med 1990;19:614–22.

*(26)* Stevens RG, Davis S, Thomas DB, Anderson LE, Wilson BW. Electric power, pineal function, and the risk of breast cancer. FASEB J 1992;6:853–60.

*(27)* Vena JE, Graham S, Hellmann R, Swanson M, Brasure J. Use of electric blankets and risk of postmenopausal breast cancer. Am J Epidemiol 1991;134:180–5.

*(28)* Wolff MS, Toniolo PG, Lee EW, Rivera M, Dubin N. Blood levels of organochlorine residues and risk of breast cancer. J Natl Cancer Inst 1993;85:648–52.

*(29)* Krieger N, Wolff MS, Hiatt RA, Rivera M, Vogelman J, Orentreich N. Breast cancer and serum organochlorines: a prospective study among white, black and Asian women. J Natl Cancer Inst 1994;86:589–99.

*(30)* Davis DL, Bradlow HL, Wolff M, Woodruff T, Hoel DG, Anton-Culver H. Medical hypothesis: xenoestrogens as preventable causes of breast cancer. Environ Health Perspect 1993; 101:372–7.

*(31)* Robbins AS, Brescianini S, Kelsey JL. Regional differences in known risk factors and the higher incidence of breast cancer in San Francisco. J Natl Cancer Inst 1997:89:960–5.


## Notes

[1]*Editor's note:* SEER is a set of geographically defined, population-based central tumor registries in the United States, operated by local nonprofit organizations under contract to the National Cancer Institute (NCI). Each registry annually submits its cases to the NCI on a computer tape. These computer tapes are then edited by the NCI and made available for analysis.


Supported in part by Public Health Service grants CA/ES62984 (National Cancer Institute and the National Institute of Environmental Health Sciences) and Institutional National Service Award 5T32CA09001 (F. Laden) and grant CA40356 (National Cancer Institute), National Institutes of Health, Department of Health and Human Services; and by American Cancer Society Faculty Research Award FRA-455 (D. J. Hunter).

We thank the registered nurses who participated in this study, Gary Chase, Karen Corsano, and Barbara Egan.

Manuscript received March 10, 1997; revised June 16, 1997; accepted July 17, 1997.