

## Population-density and county-level variation in breast cancer mortality rates among white women residing in the Northeastern and Southern United States

Susan R. Sturgeon<sup>1,\*</sup>, Barry I. Graubard<sup>2</sup>, Catherine Schairer<sup>2</sup>, Mary McAdams<sup>3</sup>, Robert N. Hoover<sup>2</sup> & Mitchell H. Gail<sup>2</sup>

<sup>1</sup>School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA 01003, USA; <sup>2</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD 20852, USA; <sup>3</sup>Information Management Services, Inc., Silver Spring, MD, USA

Received 18 February 2003; accepted in revised form 31 July 2003

**Key words:** breast cancer, mortality, United States.

### Abstract

**Objective:** We assessed the contribution of variation in risk factor prevalence to population-density and county-level variation in breast cancer mortality rates.

**Methods:** In 1995 we collected risk factor information in a telephone interview of a random digit dialed sample of: (1) 1241 women from counties in the upper and lower tertiles of population density as of 1970 in the Northeast and South of the United States (Design A); (2) 2492 women from counties in the upper and lower tertiles of 1970–1979 breast cancer mortality rates in the four populations from Design A, and; (3) 276 women in Nassau County in New York State. We calculated 1990–94 mortality ratios (MRs) adjusted for breast cancer risk factors.

**Results:** The high/low population-density fully-adjusted MRs in women  $\geq 55$  years were 1.01 (95% CI 0.9–1.2) and 1.00 (95% CI 0.8–1.2). The fully-adjusted MRs for high *versus* low mortality counties ranged from 0.95 (95% CI 0.8–1.2) to 1.29 (95% CI 1.0–1.6) in women  $\geq 55$  years.

**Conclusions:** Differences in risk factor prevalence explained higher rates in high-density *versus* low-density areas in older women. Modest elevations in the adjusted high/low breast cancer MRs among older women in certain groups of counties may reflect unidentified risk factors but more likely are due to chance.

### Introduction

Breast cancer mortality rates vary among white women across in the Northeast, South, Midwest and Western regions of the United States. For several decades, mortality from breast cancer has been higher in the Northeast than in other regions, particularly the South [1]. Breast cancer mortality rates also tend to vary by population density, with rates generally found to be higher in urban than rural areas [2, 3]. Even within the same region of the country, mortality rates can vary between counties that have similar population densities

[3]. Although these variations in breast cancer mortality rates are well recognized, there are uncertainties about the reasons for them. Our previous analysis using data from the National Health Interview Survey suggested that much of the variation across the four regions of the country could be explained by regional differences in the proportion of women who had established breast cancer risk and prognostic factors [4]. For example, we found that women in the Northeast were more likely to delay childbearing than women in the South. In the present study, we expanded upon our previous work by assessing the potential contribution of differences in the prevalence of breast cancer risk factors and prognostic factors to population-density and county-level variation in the 1990–1994 breast cancer mortality rates among white women residing in the Northeast and South.

\* Address correspondence to: Department of Biostatistics and Epidemiology, University of Massachusetts, Arnold House 407, 715 North Pleasant Street, Amherst, MA 01003, USA. Ph.: +1-413-577-1364; Fax: +1-413-545-1645; E-mail: ssturgeon@schoolph.umass.edu

## Materials and methods

### *Description of study design for the telephone survey*

Between December 1994 and April 1995, we conducted a telephone survey of white women, age 25 and older residing in selected counties from 24 states in the Northeastern and Southern regions of the United States (Table 1). To permit an analysis of variation in breast cancer mortality rates according to population density (Design A), counties in the Northeast and South were first grouped into tertiles according to their 1970 population density (number of residents per square mile), with tertile cutpoints determined by the distribution of population density in counties in the Northeast. In a further comparison of 1970 population density to 1980 population density, the 57 counties that moved from one tertile to another over the decade were omitted from the sampling frame. Counties in the intermediate tertile of population density were then excluded from the sampling frame. Thus, Design A (Table 2) included four populations of individuals; those from counties in the lowest tertile of population density in the Northeast and South (designated as low population-density) and those from the counties in the highest tertile of population density in the Northeast and South (designated as high population-density).

To permit an analysis of county-level variations in breast cancer mortality rates among the four populations included in Design A, counties within each population density stratum in Design A were further classified into tertiles according to the 1970–1979 breast cancer mortality rates in white women age-adjusted to the 1970 US population (Design B; Table 3). Tertile cutpoints were determined by the distribution of breast cancer mortality rates in counties in the Northeast. Counties in the intermediate tertile of breast cancer mortality rate were then excluded, so that Design B consisted of eight populations of individuals from four groups of counties (high- and low-mortality with high population-density; high- and low-mortality with low population-density) in both the Northeast and South. Because of particular interest in understanding the apparent high rates of breast cancer in Nassau County in New York State, it was sampled separately (Design C) even though it was also included in those counties sampled for Designs A and B.

We used a Donnelly directory of residential telephone numbers to define a population-based random digit sampling frame of telephone numbers covering each of the 12 populations above and Nassau County in New York State [5]. The samples for Design A and Design B overlapped in that individuals sampled for Design A who were from one of the populations covered by

Table 1. Definitions of Northeast and South

Northeast	Connecticut, Massachusetts, New York, Pennsylvania, Rhode Island, District of Columbia, New Jersey, Maryland, New Hampshire, Maine, Vermont, Delaware
South	Louisiana, Georgia, Kentucky, North Carolina, South Carolina, Texas, Mississippi, Arkansas, Alabama, Oklahoma, West Virginia, Tennessee

The definition of Northeast is expanded from the US census definition to include the corridor states of Delaware and Maryland and the District of Columbia.

The definition of the South excludes Virginia because of large temporal changes in county borders. Florida is also excluded because of substantial population in-migration from the Northeast.

Table 2. Number of eligible counties in the four populations of Study Design A

	Region	
	Northeast	South
Low density <sup>a,b</sup>	80	880
High density <sup>c,d</sup>	77	55

<sup>a</sup> Population density  $\leq 71$  residents per square mile (1970).

<sup>b</sup> Mean population density was 40.8 and 27.9 residents per square mile (1970) in the low population-density Northeast and South, respectively. Median population density was 43.5 and 26.0 residents per square mile (1970) in the low population-density Northeast and South, respectively.

<sup>c</sup> Population density  $\geq 246$  residents per square mile (1970).

<sup>d</sup> Mean population density was 2615.5 and 634.0 residents per square mile (1970) in the high population-density Northeast and South, respectively. Median population density was 790.0 and 530.0 residents per square mile (1970) in the high population-density Northeast and South, respectively.

Table 3. Number of eligible counties in the eight populations of Study Design B

	Region	
	Northeast	South
Low density <sup>a</sup>		
Low BC rate <sup>b</sup>	36 (14.9–26.8)	774 (0.0–26.8)
High BC rate <sup>c</sup>	20 (30.1–46.1)	56 (30.1–55.3)
High density <sup>d</sup>		
Low BC rate <sup>b</sup>	6 (24.8–26.8)	47 (17.4–26.6)
High BC rate <sup>c</sup>	43 (30.1–36.2)	1 (31.2)

Range for age-adjusted 1970–1979 white female breast cancer mortality rates.

<sup>a</sup> Population density  $\leq 71$  persons per square mile (1970).

<sup>b</sup> Age-adjusted 1970–1979 breast cancer mortality rate for white females  $\leq 26.8/100,000$ .

<sup>c</sup> Age-adjusted 1970–1979 breast cancer mortality rate for white females  $\geq 30.0/100,000$ .

<sup>d</sup> Population density  $\geq 246$  persons per square mile (1970).

Design B were included in the Design B sample. In addition, samples were drawn from six of the eight populations covered in Design B, excluding the low-density and high-density South with low breast cancer mortality populations, in order to augment the small samples from these populations that were drawn for Design A. Random samples of eligible households were selected for each design where only households (*i.e.*, residences with any number of related individuals or no more than five unrelated persons living together) were eligible for inclusion in the samples. The eligibility of a sampled household was determined by a telephone screening interview of an adult respondent at each number drawn from the sampling frame. A random selection procedure was then used to select one designated respondent from each eligible household with probabilities chosen to oversample older women, because rates of breast cancer are substantially higher among older women. We attempted to conduct interviews with at least 300 white women from each of the 12 populations and Nassau County, with the following 10-year age allocation fractions: 25–34 (5%), 35–44 (15%), 45–54 (20%), 55–64 (25%), 65–74 (20%) and 75+ (15%). Sample weights were constructed for each respondent to weight the estimates of the risk factor/prognostic distribution for: (a) differential probabilities for selecting the respondents because of their being more than one telephone or having more than one eligible respondent in the household; and (b) adjustments to match the age-distribution in the 10 year age intervals (see above) of the sample to the approximately age distribution for breast cancer deaths.

Trained interviewers administered a detailed interview including questions on basic demographic factors, lifetime residential history, reproductive and menstrual history, use of exogenous estrogens, breast cancer screening practices, personal history of prior benign

breast biopsies and breast cancer, family history of breast cancer and alcohol consumption. The average interview length was 22.9 min.

The response rates for the telephone survey for each study design ranged from 78.9 to 82.6%. A total of 4441 interviews were completed. From the 4441 interviews, a total of 1455 subjects were further excluded for the following reasons: non-white ( $n = 837$ ), region discrepancies ( $n = 161$ ), and missing or inconsistent risk factor information ( $n = 457$ ). Thus, a total of 1241, 1478, and 267 eligible study subjects had complete interview data for Designs A, B and C, respectively. The final analytic dataset included 1241 (Design A), 2502 (Design B), and 276 (Design C) subjects; these larger figures reflect the sample size augmentation procedures described above.

#### Data analysis

Separate estimates of the regional prevalence of breast cancer risk factors and prognostic factors were derived for women less than 55 years and those 55 years or older. Prevalence data derived from the survey for each group of counties were initially weighted to adjust for multiple telephones and the number of eligible persons in the household since only one eligible respondent was selected from each sampled household. Data were re-weighted to the age distribution of each of the respective groups based on 1990 population data from the US Bureau of the Census.

Information on deaths due to breast cancer during 1990–1994 was provided by the National Center for Health Statistics, and population estimates that were used to weight the observations were based on data provided by the Census Bureau. Unadjusted 1990–1994 breast cancer mortality rates for white women were calculated for the four populations of Design A, the

eight populations of Design B, and Nassau County in New York State (Design C).

Although 1970–1979 breast cancer mortality rates were used to classify counties in the study design, we used breast cancer mortality rates for 1990–1994 in the analytic models because they correspond well with the timing of the telephone survey. In addition, the use of rates from different calendar periods allowed an assessment of the persistence of the excess rates in the earlier period over time.

To compare mortality rates in different groups of counties after adjusting for differences in the prevalence of risk and prognostic factors, we calculated the adjusted mortality ratio (MR). The adjusted MR is calculated from

$$\text{MR} = \frac{R_1 \sum_{j=1}^G (n_{2j} \text{RR}_j / n_2)}{R_2 \sum_{j=1}^G (n_{1j} \text{RR}_j / n_1)} \quad (1)$$

where the  $G$  risk groups are defined by all combinations of the risk factors,  $R_1$  is the unadjusted mortality rate of breast cancer in the comparison region,  $R_2$  is the unadjusted mortality rate of breast cancer in the reference region,  $n_1$  is total number of white women in the comparison region,  $n_2$  is total number of white women in the reference region,  $n_{1j}$  is estimated number of white women in risk group  $j$  in the comparison region,  $n_{2j}$  is estimated number of white women in risk group  $j$  in the reference region,  $\text{RR}_j$  is relative risk comparing the risk among those in risk group  $j$  with the risk among those at  $j=1$ , the referent (lowest) level of all risk factors.

The estimates  $n_{ij}$  and  $n_{2j}$  are weighted by sample weights that were described earlier. The adjusted MR can be interpreted as the unadjusted MR that would be expected if the comparison region had the same proportion of women with the specified risk factors as the reference region. Equation (1) can also be interpreted by recognizing the reciprocal of the summations as one minus the attributable risk. It follows that the MR is the ratio of the baseline rates in the comparison region to that in the reference region. An analogous calculation based on risk factor prevalence data from breast cancer cases, rather than the entire population, was used in an analysis by Dean *et al.* [6].

Confidence intervals for the MR were obtained from  $\exp[\log(\text{MR}) \pm 1.96\hat{\sigma}]$ , where  $\hat{\sigma}^2$  is the estimated variance of  $\log(\text{MR})$ . The quantity,  $\hat{\sigma}^2$  is estimated as,  $D_1^{-1} + D_2^{-1} + V_1(\sum n_{1j} \text{RR}_j / n_1)^{-2} + V_2(\sum n_{2j} \text{RR}_j / n_2)^{-2}$  where  $D_1$  and  $D_2$  are the numbers of deaths in regions 1 and 2 (reference) respectively, and where  $V_1$  and  $V_2$  are the estimated variances of the sums in the denominator and numerator of Equation (1) respectively. The quan-

ties  $V_1$  and  $V_2$  were computed using a leaving-one-out jackknife method for random groups where the observations were randomly divided into 50 groups within each stratum defined by region, population-density and mortality level [7]. The relative risks,  $\text{RR}_j$ , in Equation (1) were assumed to be known constants.

#### Relative risk estimates

The relative risk estimates for each risk factor used in the main analyses are described in Table 4. Mortality relative risks (RRs) for most recognized breast cancer risk factors are not widely available in the literature. Thus, incidence RRs were used in this analysis for risk factors that have not been convincingly demonstrated to affect survival from breast cancer, including age at first live birth, age at menarche, age at menopause and family history of breast cancer [8–14]. Incidence RRs were also used for alcohol intake and biopsy-proven benign breast disease, two factors that have not been examined in relation to survival. Mortality RRs were estimated from available data for risk factors that have been shown to affect survival from breast cancer (*i.e.*, age, body mass index, and mammography use) [15–17]. Mortality RRs were also estimated for menopausal estrogen use because it has been found to be associated with an increased risk of breast cancer incidence but an unexpected decreased risk of death from breast cancer [18]. Because of the controversial nature of this association, however, multivariate adjusted MRs are provided both with and without this variable.

Incidence RRs for age at menarche, a first-degree relative with breast cancer, alcohol intake and age at menopause were derived from multivariate analyses of the Breast Cancer Detection Demonstration Project (BCDDP) case-control study [19–21]. Incidence rates for biopsy-proven breast disease were estimated from multivariate analyses of the BCDDP Follow-up Study [22]. Incidence RRs for age at first birth were obtained from a multivariate analysis of a case-control study by MacMahon *et al.* [23]. We conducted a literature review of other breast cancer studies, and found these estimates to be similar to those observed in most studies.

Mortality RRs for age were obtained from 1990–1994 breast cancer mortality rates provided by the NCHS. Mortality RRs for mammography use were estimated from clinical studies [17], and those for body mass ( $\text{kg}/\text{m}^2$ ) were obtained from an unpublished multivariate analysis of data from the BCDDP Follow-up Study [24]. Results from a multivariate analysis of the Nurses' Health Study were used to derive mortality RRs for menopausal estrogen use [25]. Education was not considered a risk factor based on unpublished data from the

Table 4. Magnitude and source of relative risks used in the models

Risk factor	Relative risks	Source of estimates
Age (years)		(NCHS)
25-34	1.0 (ref)	
35-44	14.2	
45-54	49.9	
55-64	1.0 (ref)	
65-74	1.52	
75+	1.97	
Age at menarche (years)		(19)
≥14	1.0 (ref)	
12-13	1.1	
<12	1.2	
First-degree relative with breast cancer		(20)
No	1.0 (ref)	
Yes	2.0	
Biopsy-proven benign breast disease		(22)
No	1.0 (ref)	
Yes	1.7	
Age at first livebirth		(23)
<20	1.0 (ref)	
20-24	1.2	
25-29	1.6	
30-34	1.9	
≥35	2.4	
Nulliparous	2.0	
Body mass index (kg/m <sup>2</sup> )		(24)
<21.5	1.0 (ref)	
21.5-23.3	1.0	
23.4-26.2	1.1	
≥26.3	1.4	
Alcohol intake (g/week)		(21)
None	1.0 (ref)	
<14	1.1	
14-92	1.1	
93-182	1.3	
≥183	1.7	
Age at menopause		(19)
Premenopausal	1.7	
Natural menopause before age 45	1.2	
Bilateral oophorectomy before age 45	1.0 (ref)	
Other	1.4	
Menopausal estrogen use		(25)
Yes	1.0 (ref)	
No	1.3	
Mammogram history		(17)
Never had mammogram	1.4	
Had only routine mammograms	1.0 (ref)	
Had mammogram, at least 1 for a problem	1.3	

American Cancer Society cohort study of cancer mortality (personal communication from Dr Eugenie Calle).

Multivariate mortality RRs, (*i.e.*, RRs for combinations of risk factors),  $RR_j$ , used in Equation (1) were obtained by multiplying estimated RRs for corresponding levels of each component risk factor. This procedure is justified by the fact that most component RRs were estimated with adjustment for other important risk factors and by the assumption that interactions among risk factors are negligible.

**Results**

*Risk factor prevalence data by region*

Table 5 presents the prevalence of individual risk factors for high- and low-density areas in the Northeast and South among white women 25-54 years of age. In the South, women in high-density areas tended to slightly younger and to be leaner than women in low-density areas. Similarly, women in high-density areas were also more likely to be nulliparous, delay childbearing, and to have had a prior mammogram, but less likely to abstain from alcohol. For example, 19.0% of women in the low-density areas of the South were nulliparous compared to 30.5% in the high-density areas of South. Other risk factors, including family history of breast cancer, biopsy-proven benign breast disease, and age at menopause did not vary substantially between high- and low-density areas in the South. Differences between high- and low-density areas in the Northeast were generally similar to those in South.

Table 5 also presents comparable risk factor prevalence data for women 55 years and older. In the South, women in high-density areas compared to women in low-density areas tended to have a lower body mass index, have an earlier age at menarche, and were slightly more likely to have had a prior mammogram. Women in high-density areas were also more likely to be nulliparous, to delay childbearing, and to drink alcohol. Differences between low- and high-density areas in the Northeast were similar to those found in the South, except that the distribution of age at menarche, nulliparity, and mammogram use did not vary by population density status.

*Population-density variation (Design A)*

The age-adjusted MRs for living in a high-density area compared to a low-density area among white women less than 55 years were 1.01 (95% CI 0.8-1.3) and 1.06 (95% CI 0.8-1.4) in the South and Northeast, respectively (Table 6). MRs adjusted for all factors shown in

Table 5. Population-density differences in the prevalence (%) of factors that influence breast cancer mortality among white women aged 25–54 years and 55 years and older

Risk factor	25–54 years				55 years and older			
	South		Northeast		South		Northeast	
	Low density	High density	Low density	High density	Low density	High density	Low density	High density
Age (year)								
25–34	35.1	39.3	35.9	37.8	N/A	N/A	N/A	N/A
35–44	35.9	36.0	37.1	35.9	N/A	N/A	N/A	N/A
45–54	29.0	24.7	27.0	26.3	N/A	N/A	N/A	N/A
Age (year)								
55–64	N/A	N/A	N/A	N/A	35.6	39.2	35.2	34.6
65–74	N/A	N/A	N/A	N/A	33.8	33.4	34.0	34.4
≥75	N/A	N/A	N/A	N/A	30.6	27.5	30.8	31.1
Age at menarche (year)								
≥14	27.0	22.3	24.2	23.7	41.8	33.2	30.0	30.3
12–13	44.0	58.9	57.1	48.0	49.8	53.5	53.8	54.2
12	29.1	18.8	18.8	28.3	8.5	13.3	16.3	15.5
First-degree relative with breast cancer	9.1	5.7	6.6	5.5	15.8	11.4	10.5	12.4
Biopsy-proven benign breast disease	10.8	12.4	9.7	8.9	20.8	23.8	18.8	17.0
Age at first livebirth (year)								
<20	29.0	14.6	26.1	10.3	31.0	15.7	24.2	12.9
20–24	36.5	27.0	29.4	27.8	41.4	33.5	41.8	34.6
25–29	10.8	17.5	23.1	19.0	13.3	25.2	11.9	20.9
30–34	1.8	9.6	5.6	8.8	2.3	6.1	6.8	13.4
≥35	2.9	1.0	0.5	3.4	1.6	2.1	0.0	0.8
Nulliparous	19.0	30.5	15.2	30.7	10.4	17.4	15.3	17.4
Body mass index (kg/m <sup>2</sup> )								
<21.5	23.5	29.7	18.7	25.3	22.5	24.9	13.3	17.8
21.5–23.3	14.8	21.2	20.4	19.7	14.6	16.5	19.6	11.0
23.4–26.2	21.2	18.2	21.6	15.1	25.1	28.9	21.2	33.7
≥26.3	40.5	30.9	39.3	39.8	37.8	29.8	46.0	37.5
Alcohol intake (g/week)								
None	55.3	34.3	39.0	16.2	82.5	64.3	53.1	43.9
<14	31.7	31.3	40.0	38.5	11.5	16.3	24.0	25.5
14–92	9.9	29.7	14.1	39.1	2.2	16.7	16.5	21.6
93–182	0.7	1.5	5.2	5.0	2.2	1.2	5.5	5.5
≥182	2.4	3.2	1.7	1.3	1.6	1.5	1.0	3.5
Age at menopause								
Premenopausal	83.1	83.3	88.3	81.8	1.3	4.3	4.8	2.3
Bilateral oophorectomy before age 45	4.5	8.3	1.9	4.0	9.6	7.1	11.6	5.5
Natural menopause before age 45	1.5	0.4	0.5	4.6	9.5	11.6	15.6	10.6
Other menopause	10.9	8.1	9.3	9.6	79.5	77.1	68.0	81.7
Mammogram								
All normal	27.4	33.7	29.2	35.0	59.0	60.4	64.7	66.2
1+ abnormal	14.5	18.7	14.4	17.9	12.3	15.0	12.8	14.0
No prior mammogram	58.0	47.6	56.4	47.1	28.7	23.6	22.5	19.8
Menopausal estrogen use	15.7	15.5	6.5	7.4	46.5	51.0	34.3	31.9

Table 6. Adjusted mortality ratios for white women aged 25–54 years, and 55 years and older in higher-density areas compared to lower-density areas in the Northeast and South, 1990–1994

Risk factor	25–54 years				55 years and older			
	South		Northeast		South		Northeast	
	Low density	High density	Low density	High density	Low density	High density	Low density	High density
Age alone	1.0 (ref)	1.01	1.0 (ref)	1.06	1.0 (ref)	1.18	1.0 (ref)	1.08
Age, age at menarche	1.0 (ref)	1.01	1.0 (ref)	1.04	1.0 (ref)	1.17	1.0 (ref)	1.08
Age, first-degree relative	1.0 (ref)	1.02	1.0 (ref)	1.11	1.0 (ref)	1.21	1.0 (ref)	1.06
Age, benign breast disease	1.0 (ref)	1.03	1.0 (ref)	1.07	1.0 (ref)	1.14	1.0 (ref)	1.09
Age, age at first livebirth	1.0 (ref)	0.89	1.0 (ref)	0.94	1.0 (ref)	1.04	1.0 (ref)	0.99
Age, alcohol	1.0 (ref)	0.99	1.0 (ref)	1.03	1.0 (ref)	1.16	1.0 (ref)	1.06
Age, body mass index	1.0 (ref)	1.05	1.0 (ref)	1.10	1.0 (ref)	1.20	1.0 (ref)	1.10
Age, age at menopause	1.0 (ref)	1.03	1.0 (ref)	1.09	1.0 (ref)	1.17	1.0 (ref)	1.06
Age, mammogram	1.0 (ref)	1.03	1.0 (ref)	1.07	1.0 (ref)	1.19	1.0 (ref)	1.09
Full model <sup>a</sup>	1.0 (ref)	0.95	1.0 (ref)	1.04	1.0 (ref)	1.01	1.0 (ref)	1.00
	95% CI = 0.7–1.3		95% CI = 0.7–1.5		95% CI = 0.9–1.2		95% CI = (0.8–1.2)	

<sup>a</sup> Adjusted for all other factors in the table.

the table were 0.95 (95% CI 0.7–1.3) and 1.04 (95% CI 0.7–1.5). MRs additionally adjusted for menopausal estrogen use were 0.93 (95% CI 0.7–1.2) and 1.03 (95% CI 0.7–1.5), respectively.

Among women aged 55 and older, the age-adjusted MR associated with living in a high-density area compared to a low-density area was 1.18 (95% CI 1.1–1.3) in the South and 1.08 (95% CI 1.0–1.2) in the Northeast (Table 6). After adjustment for all risk factors shown in Table 8, the corresponding MRs were 1.01 (95% CI 0.9–1.2) and 1.00 (95% CI 0.8–1.2). Additionally adjustment for menopausal estrogen use yielded MRs of 1.01 (95% CI 0.9–1.2) and 1.00 (95% CI 0.8–1.3). Adjustment for differences in the prevalence an early age at first livebirth accounted for most of the reduction in the fully-adjusted models compared to the age-adjusted models.

County-level variation (Design B)

Among the counties classified as ‘high’ or ‘low’ mortality in 1970–1979, between 0% (high-density South, high-mortality) and 80.9% (high-density South, low-mortality) remained in the same mortality category in 1990–1994 in the eight regions defined by population density and region. For example, 30.6% of the ‘low’ mortality counties in 1970–1979 in the low-density Northeast and 66.7% of the ‘low’ mortality counties in 1970–1979 in the high-density Northeast persisted in the ‘low’ mortality category over the two times periods. Comparable percentages for those that remained ‘high’ mortality over the two time periods were 40.0 and 41.9%, respectively.

Among women less than age 55 the county-level variation in breast cancer mortality rates in 1970–1979 did not persist when 1990–1994 mortality rates from the

Table 7. Mortality ratios for counties with high versus low 1970–1979 breast cancer mortality rates

	25–54 years			55 years and older		
	1970–1979	1990–1994	1990–1994	1970–1979	1990–1994	1990–1994
	MR <sup>a</sup>	MR <sup>a</sup> (95% CI)	MR <sup>b</sup> (95% CI)	MR <sup>a</sup>	MR <sup>a</sup> (95% CI)	MR <sup>b</sup> (95% CI)
Low-density South: high versus low	1.99	0.95 (0.6–1.5)	1.00 (0.7–1.5)	2.06	0.98 (0.9–1.1)	0.95 (0.8–1.2)
High-density South: high versus low	1.34	1.12 (0.7–1.8)	1.20 (0.7–1.9)	1.31	1.23 (1.0–1.5)	1.29 (1.0–1.6)
Low-density Northeast: high versus low	1.50	1.09 (0.8–1.5)	1.04 (0.7–1.5)	1.36	1.04 (0.9–1.4)	1.19 (1.0–1.4)
High-density Northeast: high versus low	1.22	1.09 (0.8–1.6)	1.12 (0.8–1.6)	1.25	1.13 (1.0–1.2)	1.07 (0.9–1.3)

<sup>a</sup> Age-adjusted.

<sup>b</sup> Adjusted for age, age at menarche, age at first livebirth, biopsy-proven benign breast disease, first-degree relative with breast cancer, body mass index, alcohol intake, mammogram history, and type of menopause.

same group of counties were used (Table 7). Fully-adjusted estimates for 1990–1994 differed little from the age-adjusted estimates, except for a slight increase in the MR for high mortality compared to low mortality counties in the high-density South. The age-adjusted MR comparing Nassau County in New York State to Northeast high-density, low-mortality counties was 1.49 in 1970–1979 and 1.16 (95% CI 0.8–1.6) using 1990–1994 mortality rates. After adjustment for the risk factors listed in Table 7, the MR for Nassau county using 1990–1994 mortality rates was 1.03 (95% CI 0.7–1.5).

Among older women, the county-level variation was attenuated in all four populations when age-adjusted 1990–1994 MRs were examined for the group of counties that had elevated rates in 1970–1979 (Table 7). The 1990–1994 age-adjusted MRs ranged from 0.98 in the low-density South to 1.23 in the high-density South. Adjustment for identified risk factors did not substantially alter these estimates, except for the low-density Northeast, where the adjusted estimate (MR = 1.19) was somewhat higher than the age-adjusted estimate (MR = 1.04). The age-adjusted RR for Nassau County compared to high-density, low-mortality areas in the Northeast using 1970–1979 mortality rates was 1.43 and using 1990–1994 rates was 1.16 (95% CI 1.0–1.3). After adjustment for the risk factors listed in Table 9, the MR was 1.09 (95% CI 0.9–1.3).

## Discussion

We observed that higher breast cancer mortality rates in high-density areas relative to low-density areas among women 55 years and older in the Northeast and South were completely explained by the higher proportion of women with established breast cancer risk and prognostic factors in high-density areas. The most important factor was age at first livebirth. Among younger women, little or no variation in breast cancer mortality rates between low- and high-density areas were observed either before or after adjustment for established breast cancer risk and prognostic factors.

County-level variation in breast cancer mortality rates in 1970–1979 within high- and low-density areas of the Northeast and South was considerably attenuated when 1990–1994 mortality rates from the same group of counties were used. The remaining elevations in mortality in some high-density counties in the Northeast and South and low-density counties in the Northeast among older women after adjustment for the factors included in our analyses were for the most part not statistically significant and could reflect chance, incomplete control for the factors included in our analyses, or the influence

of factors that we did not consider. The same attenuation in the MRs over time was seen for Nassau County. In fact, there was no statistically significant excess mortality in Nassau County compared to low-mortality, high-density areas in the Northeast using 1990–1994 rates after adjustment for known risk factors. These results suggest that the earlier high rates reflected chance fluctuations, although the possibility that the distribution of risk factors has changed over time cannot be excluded.

Prior studies of factors associated with regional variation in breast cancer mortality rates have not specifically examined reasons for population-density differences in mortality [3, 4, 26]. Several analyses of breast cancer clusters in certain areas of the Northeast [27–29] have suggested that certain environmental exposures, *e.g.* residence near chemical facilities, may play a role in the etiology of breast cancer. Major epidemiologic studies are currently underway in both Nassau County, New York and on Cape Cod, Massachusetts [30, 31].

There are a number of potential limitations to the current analysis. Foremost, it is an ecologic analysis involving multiple data sources in which the exposure and outcome information was available for the geographic regions but not for specific individuals. It is also possible that we misspecified the RR models by relying on incidence RRs for some variables and by assuming that there were no interactions among the risk factors included in the models. However, the estimates presented use the best currently available data. Another limitation of this study is that we were unable to consider important variables that could affect breast cancer survival, such as extent of disease at diagnosis or cancer treatment. However, we were able to account for some presumed surrogates for these factors, such as mammographic screening. Another limitation is that analyses focused on high/low breast cancer mortality rates included only six counties in the high-density Northeast with low breast cancer mortality rates and only one county in the high-density South with a high breast cancer mortality rate.

Despite its limitations, this study provides unique data on the prevalence of risk factors for breast cancer incidence and mortality by population density in the United States, and addresses previously unresolved questions about the reasons for population-density differences in breast cancer mortality. Our data suggest that studies focusing on population-density differences in breast cancer mortality are unlikely to be a fruitful avenue of research for identifying novel risk factors for breast cancer. Modest elevations in the adjusted high/low breast cancer MRs among older women in the groups of counties in the high-density South and

Northeast and in the low-density Northeast may reflect variations in unidentified risk factors but more likely are due to chance. In summary, this paper further emphasizes the importance of investigating the possible contribution of regional variation in established breast cancer risk and prognostic factors to regional differences in breast cancer mortality rates before considering the role of other environmental exposures.

## References

- Devesa SS, Grauman DG, Blot WJ, Pennello G, Hoover RN, Fraumeni Jr. JF (1999) *Atlas of Cancer Mortality in the United States, 1950–94*. Washington DC: US Govt Print Off, [NIH Publ No. (NIH) 99-4564].
- Kulldorf M, Feurer EJ, Miller Ba, Freedman LS (1997) Breast cancer clusters in the northeast United States: a geographic analysis. *Am J Epidemiol* **146**: 161–170.
- Blot WJ, Fraumeni Jr JF, Stone BJ (1977) Geographic patterns of breast cancer in the United States. *J Natl Cancer Inst* **59**: 1407–1411.
- Sturgeon SR, Schairer C, Gail M, Brinton LA, Hoover RN (1995) Geographic variation in mortality among breast cancer among white women in the United States. *J Natl Cancer Inst* **87**: 1846–1853.
- Lepkowski JM (1988) Telephone Sampling Methods in the United States. In: (Groves RM, Biemer PP, Lyberg LE, Massey JT, Nichols II WL, Waksberg J, eds. *Telephone Survey Methodology*. New York: Wiley, pp. 73–98.
- Dean AG, Imrey HH, Dusich K, Hall WN (1988) Adjusting morbidity ratios in two communities using risk factor prevalences in cases. *Am J Epidemiol* **127**: 654–662.
- Wolter KM (1985) *Introduction to Variance Estimation*. New York: Springer-Verlag.
- Korzeniowski S, Dyba T (1994) Reproductive history and prognosis in patients with operable breast cancer. *Cancer* **74**: 1591–1594.
- Israeli D, Tartter PI, Brower ST, Mizrachy B, Bratton J (1994) The significance of family history for patients with carcinoma of the breast. *J Am Coll Surg* **179**: 29–32.
- Fukutomi T, Kobayashi Y, Nanasawa T, Yamamoto H, Tsuda H (1993) A clinicopathological analysis of breast cancer in patients with a family history. *Surg Today* **23**: 849–854.
- Ewertz M (1993) Breast cancer in Denmark. Incidence, risk factors, and characteristics of survival. *Acta Oncol* **32**: 595–615.
- Lees AW, Jenkins HJ, May CL, Cherian G, Lam EW, Hanson J (1989) Risk factors and 10-year breast cancer survival in northern Alberta. *Breast Cancer Res Treat* **13**: 143–151.
- Ruder AM, Moodie PF, Nelson NA, Choi NW (1988) Does family history of breast cancer improve survival among patients with breast cancer? *Am J Obstet Gynecol* **158**: 963–968.
- Wang DY, Rubens RD, Allen DS, *et al.* (1985) Influence of reproductive history on age at diagnosis of breast cancer and prognosis. *Int J Cancer* **36**: 427–432.
- Mohle-Boetani JC, Grosser S, Whittemore AS, Malec M, Kampert JB, Paffenbarger Jr RS (1988) Body size, reproductive factors, and breast cancer survival. *Prev Med* **17**: 634–642.
- Adami HO, Malke B, Holmberg L, Persson J, Stone B (1986) The relation between survival and age at diagnosis in breast cancer. *N Engl J Med* **315**: 559–563.
- Morrison A (1993) Screening for cancer of the breast. *Epidemiol Rev* **15**: 244–255.
- Brinton LA, Schairer C (1993) Estrogen replacement therapy and breast cancer risk. *Epidemiol Rev* **15**: 66–79.
- Brinton LA, Schairer C, Hoover RN, Fraumeni Jr JF (1988) Menstrual factors and risk of breast cancer. *Cancer Invest* **58**: 245–254.
- Byrne C, Brinton LA, Haile RW, Schairer C (1991) Heterogeneity of the effect of family history on breast cancer risk. *Epidemiology* **2**: 2276–2284.
- Harvey EB, Schairer C, Brinton LA, Hoover RN, Fraumeni Jr JF (1989) Alcohol consumption and breast cancer. *J Natl Cancer Inst* **78**: 657–661.
- Carter CL, Corle DK, Micozzi MS, Schatzkin A, Taylor PR (1988) A prospective study of the development of breast cancer in 16,692 women with benign breast disease. *Am J Epidemiol* **128**: 467–477.
- MacMahon B, Cole P, Lin TM, *et al.* (1970) Age at first birth and breast cancer risk. *Bull World Health Organ* **43**: 209–221.
- Schairer C, Byrne C, Keyl PM, Brinton La, Sturgeon SR, Hoover RN (1994) Menopausal estrogens and estrogen-progestin replacement therapy and risk of breast cancer (United States). *Cancer Causes Control* **5**: 491–500.
- Colditz GA, Stampfer MJ, Willet WC, *et al.* (1992) Type of postmenopausal hormone use and risk of breast cancer: 12-year follow-up from the Nurse's Health Study. *Cancer Causes Control* **3**: 43–49.
- Laden F, Spiegelman D, Neas LM, *et al.* (1997) *J Natl Cancer Inst* **89**: 1373–1378.
- Laden F, Hunter DJ (1998) Environmental risk factors and female breast cancer. *Annu Rev Public Health* **19**: 101–123.
- Lewis-Michl EL, Melius JM, Kallenbach LR, *et al.* (1996) Breast cancer risk and residence near industry or traffic in Nassau and Suffolk Counties, Long Island, New York. *Arch Environ Health* **51**: 255–265.
- Ozonoff D, Ashengrau A, Coogan P (1994) Cancer in the vicinity of a department of defense superfund site in Massachusetts. *Toxicol Ind Health* **10**: 119–141.
- Jenks S (1994) Researchers to comb Long Island for potential cancer factors. *J Natl Cancer Inst* **86**: 88–89.
- Brody JG, Rudel R, Maxwell NI, Swedis SR (1996) Mapping out a search for environmental causes of breast cancer. *Public Health Rep* **111**: 494–507.