Familial patterns of the occurrence of breast cancer were studied in a population-based case-control series of black women from the Cancer and Steroid Hormone study. The risks of breast cancer among relatives of breast cancer cases were compared to those of controls who were matched for age and locale. Using the term "probands" to indicate either case or control status, significant predictors of risk to the relatives of probands included case/control status of the proband and the number of years of education completed by the proband. Genetic segregation analysis of the case families using external risks generated from SEER data indicated that the familial aggregation was consistent with Mendelian recessive transmission of a single major gene. The use of internally estimated risks, which are much less stable than the SEER risks, no longer permitted discrimination among the major locus models examined. To avoid possible reporting bias, we also performed segregation analysis on families of probands who had completed at least 12 years of education. The results from this analysis reflected the results from the entire data.

INTRODUCTION

Although breast cancer is the leading cause of cancer mortality among American black women, few studies of its risk factors have been completed (1–5). In whites, a family history of breast cancer is a major risk factor (6–10) and the presence of genetic factors that increase susceptibility is indicated by observations of multigenerational pedigrees that include many affected women (11–13). Genetic analyses of population-based series of white breast cancer families have generally supported the presence of a dominantly expressed locus that increases susceptibility for breast cancer (14–16). However, in a population-based study of families ascertainment through premenopausal, synchronously affected, and bilaterally affected probands, segregation of a major locus with recessive expression provided a better fit to the data (17, 18). Evidence that the distribution of causes or risk factors for breast cancer in blacks may differ from those in whites has been supported by observations of a larger percentage of estrogen receptor-negative tumors in blacks (19) and a different distribution of age of onset in blacks, with higher rates among younger blacks than among whites (20, 21).

In this population-based study, we assess the association of a family history of breast cancer with risk for breast cancer in black women. We also fit genetic models to describe the familial patterns of this disease.

MATERIALS AND METHODS

Data and Methods

These data are a subset of those obtained through the Cancer and Steroid Hormone Study, conducted by the Centers for Disease Control. Details of the population and data collection have been presented (22). Briefly, women ages 20 to 54 years who had a histologically confirmed first primary cancer of the breast were ascertained between December 1, 1980, and December 31, 1982. Cases were obtained from eight population-based tumor registries (Atlanta, Connecticut, Detroit, Iowa, New Mexico, San Francisco, Seattle, and Utah) that are part of the SEER3 program of the National Cancer Institute. Overall, 4730 (80%) of the 5896 women who met the case criteria were interviewed. Of these, 490 were self-reported to be black and non-Hispanic and were retained for the present analysis. Controls, frequency matched to the age distribution of the breast cancer cases within geographic regions, were selected by random-digit dialing (23). Overall, 4754 (83%) of the 5698 eligible controls were interviewed. Of these, 485 identified themselves as black and non-Hispanic. Race-specific response rates could not be obtained from these data because race was not accessed on initial contact of probands.

Data on history of breast cancer among mothers, sisters, daughters, and second-degree female relatives were obtained from face-to-face interviews with the cases and controls. Since these individuals (the cases and controls) brought their families into the study, we shall call them "probands." Current age (or age at death) and age at onset for breast cancer were obtained for mothers, sisters, daughters, and grandmothers. Analysis was restricted to mothers and sisters only because the daughters were too young to be at appreciable risk, and reported medical histories from second-degree relatives are generally not considered very reliable (24). A relative was considered to be affected if a primary cancer of the breast was reported by the proband. Breast cancer status was considered unknown if the relative had cancer at some unknown site or the relative's cancer status was unknown. Otherwise, the relative was classified as unaffected.

Educational level of the proband was recorded as the number of self-reported years of school completed. This variable was dichotomized into higher (12 years or more of school) and lower (less than 12 years of school) levels. Self-reported income of the household of the proband was recorded in categories with $5,000 increments. For analysis, income was also dichotomized, with probands having "higher income" if their household earned at least $20,000 per annum and "lower income" if less than $20,000 per annum. Respondents were self-reported as either premenopausal, perimenopausal, postmenopausal from surgery, postmenopausal, pregnant or nursing, or unknown. Menopausal status was dichotomized, with premenopausal and nursing or pregnant women forming the referent group. To investigate the possibility that recurrence risks among relatives of younger case probands might be higher than among relatives of older case probands, we created two age strata; cases and controls were grouped according to whether they were

3 The abbreviations used are: SEER, Surveillance, Epidemiology, and End Results program; RR, relative risk; CI, confidence interval; SES, socioeconomic status.
younger or older than 50 years of age. Individuals with missing or unknown data were excluded from analyses using those variables.

We eliminated families of both cases and controls from analysis when current age was missing on all or at least two relatives (n = 23). In those families (n = 22) where only one relative had a missing current age, an estimated current age (Rc) was calculated as
\[ R_c = P_c - X_c + X_r \]

where \( P_c \) is the reported current age of the proband, \( X_c \) is the mean current age of all probands, and \( X_r \) is the mean current age of all relatives of the missing type (i.e., mothers or sisters). Relatives with unknown affection status for breast cancer (n = 20) were also eliminated. After exclusion of all families that were missing information, as indicated above, 477 control probands yielded data on 471 mothers and 1028 sisters; 475 case probands with 469 mothers and 1099 sisters remained.

For the purposes of segregation analysis, only data from the families of the case probands were used (i.e., probands plus their mothers and sisters), because the analysis conducted required ascertainment through affected individuals. Five families in which more than one sibling had an unknown affection status were eliminated. Individuals who were less than 16 years of age at the time of ascertainment were also removed from the analysis because they are not at risk for breast cancer. Segregation analysis was further restricted to the families of cases who reported having at least one sister, because data from singleton families are not informative for selecting among the various genetic hypotheses. This resulted in 369 probands and 1396 mothers and sisters of the probands on which segregation analysis was conducted.

Statistical Methods

Preliminary Analyses. Mean levels of education, income, and sibship size between case and controls were compared by using \( t \) tests. Variability of demographic variables among centers was assessed by analysis of variance. Tests for effects from each center were obtained by linear contrasts among the means; the Bonferroni inequality was used for assessing the significance of the findings (25). Among probands, the effects of income and education on risk of breast cancer were further assessed by unconditional logistic regression to evaluate evidence for a direct effect of these variables on breast cancer risk.

Epidemiological Analyses. Risks for breast cancer among sisters and mothers of probands were assessed by using Kaplan-Meier product-limit estimators (26). Cox proportional hazards models (27) were used to assess evidence that risk for breast cancer in the mothers and sisters of the proband could be predicted by the proband's (a) breast cancer status, (b) educational status, (c) income status, (d) center of residence, (e) sibship size, (f) age stratum, and (g) menopausal status. In addition, statistical interactions between both the relationship to the proband and the case status of the proband with the other possible predictors (as well as the case status by relationship interaction) were evaluated. For multivariable analyses data from both the mothers and sisters were incorporated by including an indicator for the type of relationship the individual had to the proband. Multivariable models were fitted including all of the proband-specific covariates evaluated in univariate analyses. In addition, backward selection methods were also fitted to select a parsimonious model. These models were fitted both with and without center effects. Because the outcomes among family members may be correlated, the standard errors of estimates from these analyses are likely to be slightly underestimated.

Genetic Segregation Analyses. Data from these families were further analyzed by complex segregation analysis (28, 29), to disentangle major gene effects from other sources of familial resemblance. All analyses were conducted with the computer program POINTER, which utilizes maximum likelihood methods to estimate parameters and to evaluate competing hypotheses. The most general model that was fitted has been called the "mixed" model (29). This model assumes that the population distribution of susceptibility for disease follows a mixture of normal distributions, and that an individual becomes affected if her susceptibility exceeds a threshold on the susceptibility continuum because of major genetic, polygenic, and/or environmental factors. The mixed model estimates the gene frequency (\( q \)) at a putative major gene locus affecting susceptibility, displacement (\( t \)) of the mean susceptibilities of homozygotes at the major gene locus, degree of dominance at the major gene locus (\( d \)), and non-mendelian heritability (\( H \)) of susceptibility, as well as the proportion of "sporadics" (\( x \)). Sporadics are individuals who became affected despite not having been predicted to be at excess risk from these models. Models with various restrictions were tested against the mixed model, using the likelihood ratio criterion.

Because of the sampling scheme used, it was necessary to correct for the mode of ascertainment. The ascertainment probability (\( \pi \)), which is the probability that an affected individual is a proband, was set at 0.01 because the lifetime probability of ascertainment was very low and because each family included only a single proband.

Allowance was made for the variable risk of developing breast cancer risk due to age by calculating individual liabilities (risk) from the age-specific incidence rates in black females obtained from SEER (20). Cumulative incidences were calculated to the midpoint of each interval for each liability class (30). We defined 10 liability classes based on age at onset (if affected) or age at examination (if unaff acted): 16-24 (0.00004); 25-29 (0.00034); 30-34 (0.00154); 35-39 (0.00410); 40-44 (0.00823); 45-49 (0.0142); 50-54 (0.0216); 55-64 (0.0349); 65-74 (0.0550); 75+ (0.0735). Because of the concern about underreporting of familial cases, we also used the controls (including data from sisters and mothers) to generate internally estimated risks for a separate segregation analysis. The internally estimated risks were generally lower than those from SEER, with the maximal risk of 0.0439 being reached in the 65-74-year-of-age interval. Individuals were assigned to liability classes on the basis of their age at onset if affected or on their age at examination or age at death if unaffected.

We also conducted similar analyses on the subset of the families of probands who had completed 12 years or more of school, because of the concern for possible underreporting in the sample of families in which the proband had less education. The sample studied in this analysis comprised 279 probands and their 1059 mothers and sisters.

RESULTS

Preliminary Analyses. The median age of all probands was 45 years. The current age or age at death of mothers of probands ranged from 22 to 98 years of age, with a median of 63, and the current age or age at death of sisters of probands ranged from 0 to 76 years of age, with a median of 40. In 4.4% of the control and 8.0% of the case families a first-degree relative was affected. For multivariable analyses data from both the mothers and sisters were used (i.e., probands plus their mothers and sisters) to generate internally estimated risks for a separate segregation analysis. The internally estimated risks were generally lower than those from SEER, with the maximal risk of 0.0439 being reached in the 65-74-year-of-age interval. Individuals were assigned to liability classes on the basis of their age at onset if affected or on their age at examination or age at death if unaffected.

The characteristics of the population are given in Table 1. The less educated probands were significantly older [46.6 ± 7.2 (SD) years] than more educated probands [42.3 ± 8.0 years]. After allowing for age, risk of breast cancer was not significantly predicted by either educational status of the proband (RR = 1.26, 95% CI = 0.93-1.69) or her income status (RR = 0.94, 95% CI = 0.79-1.26). The Spearman correlation between educational level and income was 0.37 (\( P < 0.001 \)).

The mean educational level of the probands varied among the centers. On the average, probands from San Francisco had completed 1.30 more years of education than probands from other areas, but no other significant differences among centers were noted after correction for multiple testing.

Epidemiological Analyses. Cumulative risk of breast cancer to age 65 years reached 2.1% ± 0.8% (SD) in control mothers and 7.6% ± 3.7% in control sisters (Table 2). For control mothers, the lifetime risk is considerably less than would be expected from both SEER registry data (20) and comparison with reports of risk for breast cancer among mothers of white controls from the same study (10). It is notable that among the mothers and sisters of control probands (n = 134, 315, respec-

1794

Downloaded from cancerres.aacrjournals.org on July 20, 2017. © 1991 American Association for Cancer Research.
mothers and sisters of controls who had completed at least 12 years of education. For the more educated probands, the risk to age 54 among the 187 control mothers was 2.0%, and among the 371 sisters it was 3.1% (Table 3). These risks are similar to the 2.6% cumulative rate to age 54 from SEER.

Table 4 presents results from fitting Cox-proportional hazards models to predict risk for breast cancer among mothers and sisters of the probands. No significant effect was observed for center of data collection on risk for breast cancer among relatives when the effects due to each center were jointly tested ($\chi^2_{6} = 12.18, P = 0.09$). However, from the univariate analyses shown in Table 4, risks were higher for mothers of probands from San Francisco.

Results from univariate Cox-proportional hazards modeling are shown in Table 4. Among the mothers, case status of the proband was a predictor of breast cancer risk, as were higher income level and age of the proband. Among sisters of probands, higher educational level of the proband was a predictor of breast cancer risk. After allowance for the proband’s educational level, the relative risk among mothers of breast cancer probands versus that among mothers of control probands was 2.13 (95% CI = 1.01–4.49), while among sisters the risk, allowing for educational level, was 1.55 (95% CI = 0.73–3.29). Similarly, after allowance for income level, the relative risk for breast cancer among mothers of breast cancer probands versus the risk among mothers of control probands was 2.28 (95% CI = 1.04–4.97), while among sisters, the risk after allowing for income level was 1.70 (95% CI = 0.78–3.73).

Results from multivariable analysis are given in Table 5. Higher educational level and case status of the proband were significant predictors of risk, with the sisters of probands also being at significantly higher risk than mothers of probands. Implementation of backward selection procedures, including all of the covariates indicated in Table 5 as well as center effects, led to a model with effects from the educational status (RR = 1.55, 95% CI = 0.73–3.29), case status of the proband (RR = 2.13, 95% CI = 1.01–4.49), residence in San Francisco versus other centers (RR = 2.28, 95% CI = 1.04–4.97), while among sisters, the risk after allowing for income level was 1.70 (95% CI = 0.78–3.73).

### Table 1 Distribution of characteristics among cases and controls

<table>
<thead>
<tr>
<th>Cases</th>
<th>Mean ± SD</th>
<th>n</th>
<th>Controls</th>
<th>Mean ± SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (yrs)</td>
<td>43.1 ± 7.6</td>
<td>(475)</td>
<td></td>
<td>43.8 ± 8.4</td>
<td>(477)</td>
</tr>
<tr>
<td>Education &lt;12 yr</td>
<td>46.3 ± 6.9</td>
<td>(113)</td>
<td></td>
<td>46.9 ± 7.6</td>
<td>(137)</td>
</tr>
<tr>
<td>≥12 yr</td>
<td>42.1 ± 7.5</td>
<td>(361)</td>
<td></td>
<td>42.6 ± 8.9</td>
<td>(338)</td>
</tr>
<tr>
<td>Income ($) &lt;20,000</td>
<td>43.4 ± 7.9</td>
<td>(259)</td>
<td></td>
<td>43.7 ± 8.7</td>
<td>(272)</td>
</tr>
<tr>
<td>≥20,000</td>
<td>42.3 ± 7.0</td>
<td>(187)</td>
<td></td>
<td>43.8 ± 7.7</td>
<td>(177)</td>
</tr>
<tr>
<td>No. of sisters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (yr)</td>
<td>2.4 ± 2.0</td>
<td>(475)</td>
<td></td>
<td>2.2 ± 2.0</td>
<td>(477)</td>
</tr>
<tr>
<td>Education &lt;12 yr</td>
<td>2.5 ± 1.8</td>
<td>(113)</td>
<td></td>
<td>2.4 ± 2.2</td>
<td>(137)</td>
</tr>
<tr>
<td>≥12 yr</td>
<td>2.4 ± 2.0</td>
<td>(361)</td>
<td></td>
<td>2.1 ± 1.9</td>
<td>(338)</td>
</tr>
<tr>
<td>Income ($) &lt;20,000</td>
<td>2.4 ± 1.9</td>
<td>(259)</td>
<td></td>
<td>2.3 ± 2.0</td>
<td>(272)</td>
</tr>
<tr>
<td>≥20,000</td>
<td>2.3 ± 2.0</td>
<td>(187)</td>
<td></td>
<td>2.1 ± 1.9</td>
<td>(177)</td>
</tr>
<tr>
<td>Age of mother (yr)</td>
<td>62.4 ± 11.6</td>
<td>(469)</td>
<td></td>
<td>62.2 ± 13.0</td>
<td>(471)</td>
</tr>
<tr>
<td>Age of sisters (yr)</td>
<td>35.9 ± 13.6</td>
<td>(1,099)</td>
<td></td>
<td>39.6 ± 14.8</td>
<td>(1,028)</td>
</tr>
<tr>
<td>Education (yr)</td>
<td>12.9 ± 3.0</td>
<td>(474)</td>
<td></td>
<td>12.4 ± 2.7</td>
<td>(469)</td>
</tr>
<tr>
<td>Income ($)</td>
<td>21,456 ± 13,984</td>
<td>(443)</td>
<td></td>
<td>21,535 ± 13,891</td>
<td>(444)</td>
</tr>
</tbody>
</table>

* Mean ages of less educated probands (cases and controls) are greater than that of more educated probands ($P < 0.001$).

* Cases are more educated than controls ($P < 0.05$).

* Effects due to other centers not shown.
Segregation analysis using all the families ascertained through a breast cancer case revealed that an autosomal recessive model was adequate to explain the familial clustering observed (Table 6). When estimation of parameters for the full model was attempted, the parameters \(d, H, \) and \(x\) became fixed at 0.0, indicating that they were superfluous in describing the familial patterns of disease. Very similar results were obtained when only families in which the proband had completed 12 or more years of education were analyzed (results not shown). Segregation analysis was also performed on the internally estimated risks generated from the controls. The results from this analysis were similar to those obtained by using SEER risks, but it was no longer possible to discriminate between major locus models. For these data, the autosomal recessive model was the most parsimonious, but the likelihoods for the additive and dominant major locus models were not very different from the recessive model, with no significant differences among the major gene models (results not shown). We also analyzed a subset of the data restricting the sample to premenopausal breast cancer cases only. Although models for no transmission and multifactorial inheritance alone could be rejected, discrimination between the major locus models was not possible (results not shown).

**DISCUSSION**

In this study, we assessed family history as a predictor for risk of breast cancer in a population-based study of American black women. Studies incorporating other covariables that affect risk for breast cancer are reported elsewhere (31). As is also seen among whites, we find consistent evidence that family history of breast cancer in a first-degree relative is predictive of an individual's breast cancer risk (6–10), and the relative risks are of the same order for blacks as they are for whites (7, 10).

We also found a significant effect of educational level and income of the proband on the risk for breast cancer in her relatives. This effect may arise for one of three reasons. First, the effect may reflect a bias, with the more educated probands more accurately reporting breast cancer occurrence in their relatives. Second, the SES of the probands (which is partially determined by income and education) may affect their risk of developing or being diagnosed as having breast cancer. In the current study, we found that the case probands had finished somewhat more years of school than control probands but did not observe any significant differences in income between case and control probands. It seems unlikely, therefore, that SES should have an effect on risk for breast cancer for relatives when a major effect of SES on risk for breast cancer is not seen among the probands. In a prior study, however, blacks from working-class census tracts were found to have a later age at onset of breast cancer and to have a lower incidence of breast cancer than blacks from census tracts in which the woman had a higher SES (5). Devesa and Diamond (1) also found that among blacks, educational level and income were both predictive of risk for breast cancer. Finally, SES measures could act as surrogate measures for age or center effects. In multivariable analysis, however, educational level was observed to have an effect independent from that of center, and was a better predictor of breast cancer risk than the age of the proband.

In this study, we cannot discriminate well between the possibility that the educational level of the proband reflects the accuracy of her reporting of breast cancer in her relatives, versus the possibility of a direct effect of the SES of the proband on risk for breast cancer in her relatives. We therefore performed segregation analysis twice, once using all of the data, and then restricting the sample to the more educated probands. The results from the segregation analysis on the total sample indicated that a recessive major gene alone was consistent with the distribution of breast cancer in blacks. For the analysis restricted to the smaller number of more educated probands, it was not possible to discriminate among the major locus models examined, but the best-fitting model was very similar to that from the entire set of data.

The interpretation of results from these segregation analyses must include several concerns. The number of families that were informative for segregation analysis was small, and the asymptotic theory upon which the hypothesis tests were constructed may not apply (32–34). Segregation analysis also has inherent limitations. The segregation pattern may mimic mendelian genetic transmission when in fact other processes may result in disease transmission in families. Methods for allowing for the variable age at onset of breast cancer are still being developed (35). Evidence of more than a single major genetic effect is difficult to evaluate by segregation analysis. Finally, nuclear families provide limited information for differentiating among major gene hypotheses or for comparing major genetic to environmental or more complicated models, because the segregation patterns of disease through several generations cannot be observed. Furthermore, measured environmental data were not available for the relatives of probands. Our inability to allow for environmental risk factors for breast cancer may bias the results of this study, particularly if any unmeasured...
environmental and life-style factors are strongly correlated among family members. Evidence for major gene effects can also be obtained from genetic linkage analysis; this approach has recently provided evidence for a major gene on chromosome 17q12–21, which greatly increases susceptibility for breast cancer (36). Genetic linkage studies in blacks may help to resolve any differences in etiology for breast cancer among women from high-risk pedigrees.

We found evidence that measures of SES affect risk of reported breast cancer in the relatives of the probands. This effect may reflect a bias in the knowledge or reporting by the probands of breast cancer among their relatives. The risk of breast cancer among relatives of the less well-educated probands was lower than it was for relatives of the more educated probands for both the case as well as the control probands. Thus, although the less well-educated probands may underreport breast cancer among their relatives, this underreporting should not greatly alter the estimate of the family history effect. Results from segregation analysis were consistent with the segregation of a major gene affecting susceptibility for breast cancer, whether or not allowance was made for the varying educational levels of the probands. Results from segregation analysis of all of the data using SEER risks were most consistent with segregation analysis were limited by the small sample size and by having only covariate information for the probands. Thus, we thank Dr. Sherri Bale and Phyllis Wingo for their assistance.

REFERENCES

Familiality of Breast Cancer and Socioeconomic Status in Blacks

Christopher I. Amos, Alisa M. Goldstein and Emily L. Harris


Updated version

Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/51/7/1793

E-mail alerts

Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions

To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.