

# Detecting an Association between Socioeconomic Status and Late Stage Breast Cancer Using Spatial Analysis and Area-Based Measures

Jill Amlong MacKinnon,<sup>1</sup> Robert C. Duncan,<sup>1,2</sup> Youjie Huang,<sup>3</sup> David J. Lee,<sup>1,2</sup> Lora E. Fleming,<sup>1,2</sup> Lydia Voti,<sup>1</sup> Mark Rudolph,<sup>1</sup> and James D. Wilkinson<sup>2</sup>

<sup>1</sup>Florida Cancer Data System, Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine;

<sup>2</sup>Department of Epidemiology and Public Health, Miami, Florida; and <sup>3</sup>Florida Department of Health, Bureau of Epidemiology, Tallahassee, Florida

## Abstract

**Objectives:** To assess the relationship between socioeconomic status (SES) and late stage breast cancer using the cluster detection software SaTScan and U.S. census-derived area-based socioeconomic measures.

**Materials and Methods:** Florida's 18,683 women diagnosed with late stage breast cancer (regional or distant stage) between 1998 and 2002 as identified by Florida's population-based, statewide, incidence registry were analyzed by SaTScan to identify areas of higher-than-expected incidence. The relationship between SES and late stage breast cancer was assessed at the neighborhood (block group) level by combining the SaTScan results with area-based SES data.

**Results:** SaTScan identified 767 of Florida's 9,112 block groups that had higher-than-expected incidence of late stage breast cancer. After controlling for patient level insurance status, county level mammography prevalence,

and urban/rural residence in the logistic regression model, women living in neighborhoods of severe and near poverty were respectively 3.0 and 1.6 times more likely to live in areas of higher-than-expected incidence of late stage breast cancer when compared with women living in nonpoverty. Additionally, areas in the lowest quartile of mammography usage were almost seven times more likely to have higher-than-expected incidence than areas in the higher quartiles.

**Conclusions:** In addition to confirming the importance of mammography, results from the present study suggest that "where" you live plays an important role in defining the risk of presenting with late stage breast cancer. Additional research is urgently needed to understand this risk and to leverage the strengths and resources present in all communities to lower the late stage breast cancer burden. (Cancer Epidemiol Biomarkers Prev 2007;16(4):756-62)

## Introduction

One of the fundamental principles of public health is that socioeconomic status (SES) is a strong determinant of health. This strong relationship has been documented for centuries, dating back to ancient Greece, Egypt, and China (1-3). More and more health researchers believe that a narrow focus on individuals outside of historical, social, and biophysical contexts limits the understanding of disease etiology, health, and intervention modes.

Breast carcinoma is the most common cancer in women in the United States, particularly in Florida, and the second leading cause of cancer death in women (4). Approximately 64,000 Florida women were diagnosed with a primary breast cancer between 1998 and 2002.

The relationship between SES and breast cancer is complex. Overall incidence rates increase with increasing SES (5). However, the incidence of distant breast cancer is higher among women of low SES (6). This disproportionate distribution is consistent with reports indicating that women of lower SES have lower rates of mammography screening and lower

adherence to recommended follow-up after abnormal mammograms (6).

Racial differences have been documented in breast carcinoma incidence, mortality, and survival rates. Although breast cancer incidence is somewhat lower among African American women than among White women in the United States, breast cancer mortality is consistently higher among African Americans (7, 8). These racial disparities in breast carcinoma mortality and survival rates can be partially explained by cancer stage distribution at the time of diagnosis, treatment quality, treatment adherence, etc., which may in turn be related to SES.

In addition to race, social and economic factors have been associated with cancer mortality and survival. Baquet et al. found that SES, more than race, predicts the likelihood of a group's access to education, certain occupations, and health insurance, as well as income level and living conditions, all of which are associated with a person's chance of being diagnosed with late stage disease and surviving cancer (9).

The lack of readily available patient level SES information in surveillance systems limits large scale research efforts to fully study this phenomenon, subsequently limiting our total understanding of the contributions of SES to breast cancer incidence and mortality. Despite growing recognition of the magnitude of SES inequalities in health, very few, if any, SES data are available in most U.S. public health surveillance systems (10, 11).

One solution to the absence of SES data in surveillance systems gaining increased recognition involves geocoding and the use of area-based SES measures to characterize persons in both public health data bases and total population (12). These census-derived area-based measures can be conceptualized as

Received 5/16/06; revised 1/19/07; accepted 1/31/07.

**Grant support:** Florida Department of Health contract COANF, Centers for Disease Control through the National Program of Cancer Registries cooperative agreement U55/CA21918-04, University of Miami Miller School of Medicine, James and Esther King Biomedical Research Program of Florida Department of Health, and Sylvester Comprehensive Cancer Center.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**Requests for reprints:** Jill A. MacKinnon, Florida Cancer Data System, University of Miami Miller School of Medicine, P.O. Box 016960 (D4-11), Miami, FL 33101. Phone: 305-243-3426; Fax: 305-243-4871. E-mail: jill\_mackinnon@miami.edu

Copyright © 2007 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-06-0392

meaningful indicators of SES context in their own right and not merely “proxies” for individual level data. These measures provide information on not only the area’s residents (i.e., its composition) but also area level characteristics not reducible to the individual level (e.g., concentration of poverty, absence of a nearby clinic, adjacency of toxic waste site, environmental attributes, etc.; ref. 13).

Assignment of SES from ecological data can be challenging. The literature is replete with different SES indices that are composed of various census variables. Krieger et al. suggest that efforts to monitor the U.S. socioeconomic inequalities in health using area-based measures will be best served by those tract or block group measures that are (a) most attuned to capturing economic deprivation, (b) meaningful across regions and over time, and (c) easily understood and hence based on readily interpretable variables with *a priori* categorical cut points. The U.S. census variable “percentage of persons living below the U.S. poverty line” meets all of these criteria (14).

This research uses only late stage disease because late stage disease is most often implicated with lower SES (15, 16). Using cancer surveillance data, this study incorporated spatial statistical methodology and U.S. census-derived area-based socioeconomic measures as an initial evaluation of the degree to which SES is associated with the incidence of late stage breast cancer in Florida. This research shows that the lack of individual SES data need not limit research efforts directed at studying the effects of SES on disease. Additionally, the methods used in this research can “map” geographically defined subpopulations of excess disease, allowing cancer control professionals to target interventions. Furthermore, these geographically defined areas can serve as a baseline for evaluating the effectiveness of future interventions.

## Materials and Methods

**Overview.** This was a cross-sectional study of Florida women diagnosed with late stage (regional and distant stage) breast cancer between 1998 and 2002, as identified by the State of Florida incident cancer registry, the Florida Cancer Data System (FCDS). A woman’s SES was based from the census variable “ratio of income to poverty” of her block group of residence at diagnosis (obtained from the individual patient data). Block group was chosen because it is the smallest geographic unit for which U.S. census data are available and allows for a more precise SES assignment than either census tracts or zip codes (which can change over time and are more likely to contain heterogeneous populations due to their larger size). Of note, the words “neighborhoods” and “area” used throughout this paper are intended to be synonymous with the term “block group(s).”

The unit of analysis for this study was the block group. As depicted in Fig. 1 and discussed in detail below, the patient level data from Florida’s statewide, population-based registry were aggregated by race and age group at the block group level. The aggregated patient data served as the numerators and the aggregated census population (also stratified by race and age group) as the denominators for the spatial analysis. The results of the spatial analysis identified block groups of higher-than-expected late stage breast cancer incidence. Logistic regression analysis was used to model the probability of a block group having a higher-than-expected incidence of late stage breast cancer as a function of the block group SES. The dependent variable was the incidence of late stage breast cancer (i.e., block groups categorized as having either an excess or expected incidence by the cluster detection software described below). The independent variables included area-based measures: (a) census-derived SES, (b) late stage breast

cancer patient-derived insurance status, (c) county level urban/rural residence, and (d) county level mammography prevalence.

**Breast Cancer Data.** The study was conducted using data collected for female breast cancer patients, residents of Florida, diagnosed, from January 1, 1998 through December 31, 2002, with late stage breast cancer disease (regional or distant) by FCDS. FCDS is the legislatively mandated, population-based central cancer registry for the State of Florida.

Cases in the FCDS are abstracted from patient medical records by hospitals, freestanding ambulatory surgical facilities, radiation therapy facilities, private physicians, and death certificates and contain basic demographic, cancer diagnosis (primary site and cancer histology), tumor staging, and treatment information. External estimates designate FCDS completeness percentage to be >95% (as determined by external quality control audits). FCDS is part of the Centers for Disease Control National Program of Cancer Registries and is nationally certified by the North American Association of Central Cancer Registries at its highest level, “gold certification,” for meeting/exceeding the standards for completeness, timeliness, and quality.

Data collected and coded by FCDS are in accordance with national standards as put forth by North American Association of Central Cancer Registries. *International Classification of Diseases for Oncology, 3rd edition* was used to code primary site and morphology. During the study time period, tumors were staged using the SEER General Summary Stage, which categorizes stage of disease as *in-situ*, local, regional, or distant. For the purpose of this analysis, only cases coded as regional or distant stage at diagnosis were included in the study. Regional and distant stage tumors were combined in one “late stage” category (of note, because regional and late stage data were combined into one category of late stage, the revised criteria to the summary stage schema in 2001 had no effect on this analysis).

The file used for this analysis was a de-identified file containing over two million patient level cancer records diagnosed between 1981 and 2002. Although there were no patient identifiers on the file, the file contained geocoded information for each case (census tract, block group, block, longitude, and latitude), geocoded to the 2000 U.S. census by a contract vendor. This study was approved by the University of Miami Institutional Review Board.

The criteria for selecting unduplicated primary breast cancer cases included *International Classification of Diseases for Oncology, 3rd edition* primary site codes C50.0 through C50.9, Florida resident at time of diagnosis, diagnosed between 1998 and 2002, and stage of disease at diagnosis recorded as regional or distant. The dataset used for this analysis was unduplicated by FCDS.

There were 19,023 late stage female breast cancers recorded among Florida residents for the study period. A total of 340

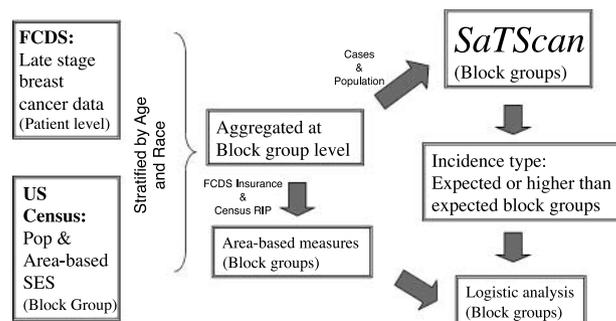


Figure 1. Data structure and flow.

cases were eliminated from the extracted data set because the records contained an "other/unknown" race code ( $n = 31$ , 0.2%) or were not assigned to a block group ( $n = 309$ , 1.6%). The final analytic data set contained 18,683 cases.

### Census Data

**Population.** Block group level population was downloaded from the U.S. 2000 Census Bureau Web site. The population data were stratified at the block group level by gender, race, and age group. Because of difficulties obtaining mutually exclusive race/ethnicity data from the U.S. census at the level of specificity necessary for the spatial software, the analysis was done using race only, without regard to Hispanic ethnicity.

There are 9,112 block groups in Florida. The individual block group annual 2000 population estimates were multiplied by five to provide 5-year sex/race/age-specific denominators for the 1998 to 2002 study period.

### Area-Based Measures

**Socioeconomic status.** For this project, the census variable ratio of income to poverty was used for the assignment of SES. The characteristics of the family used to determine poverty status are the number of people in the household, the number of related children under 18, and whether the primary householder is over age 65. An income threshold is determined given a particular family's set of characteristics; if that family's income is below that threshold, the family is considered to be in poverty. The household income divided by this threshold (as established by the U.S. government) is called the ratio of income to poverty. By way of example, the 2000 poverty threshold for a family of four was \$17,463 (17).

The U.S. census provides the number of persons within each of the nine ratio of income to poverty categories (<0.50, 0.50-0.74, 0.75-0.99, 1.0-1.24, 1.25-1.49, 1.50-1.74, 1.75-1.84, 1.85-1.99, and >2.0) at the block group level (U.S. census variables, P088002-P088010). The category specific percentages were computed in SPSS.

Based on methodology by Krieger et al., the nine ratio of income to poverty categories were recoded into three categories and classified as "severe poverty" (ratio, <0.99), "near poverty" (ratio, between 1.00 and 1.99), and "non-poverty" (ratio, >1.99; ref. 18). For each block group, the percentage of residents in severe poverty, near poverty, and nonpoverty were calculated (note these percentages add to 100%). The total block group was then assigned the SES description of severe poverty, near poverty, or nonpoverty according to the largest value among the three categories. There were 16 block groups for which two of the ratio of income to poverty categories were identical. For these 16 block groups, SES was assigned by a flip of a coin.

**Insurance.** Because neither the U.S. census nor any other population-based Florida survey contains block group level insurance data, the area-based measure for "insurance status" was obtained from the FCDS database. The woman's insurance status at the time of late stage cancer diagnosis in the FCDS file was recoded from 17 different values into four mutually exclusive categories: uninsured, private, medicare, and medicaid. The insurance information obtained from the FCDS data was aggregated by block group, and the proportion for each of the four values was calculated. The block group insurance status was assigned based on plurality as described for SES.

**Urbanization.** Urbanization was based on the county of residence at diagnosis. Urban/rural continuum codes (Beale codes; ref. 19) were used to characterize counties based on their population size, degree of urbanization, and proximity to a large metropolitan area. Beale codes were recoded into two categories: urban counties and rural counties. Block groups were classified based on the county in which the block group was located.

**Mammography Prevalence.** Mammography usage was based on the 2002 county level data collected by the Behavioral Risk Factor Surveillance System for the State of Florida. The race specific prevalence was not complete for each county. Therefore, individual county level prevalence data for all races were recoded into quartiles. The block group was assigned the quartile computed for the county in which the block group was located.

**Spatial Analysis Software.** SaTScan version 5.0 was used to identify areas within Florida that had an excess of late stage breast cancer. The spatial scan statistic used by SaTScan is a cluster detection test that is able to detect both the location of clusters and evaluate their statistical significance. The SaTScan software can analyze spatial, temporal, and space-time data using the spatial, temporal, or space-time scan statistics (20, 21). For this study, only the spatial analysis was done.

Properties that make the spatial scan statistic suitable for geographic surveillance include (a) its ability to take into account the uneven geographic distribution of cases and population densities, (b) its lack of assumptions about cluster size or location, (c) its ability to adjust for multiple testing (a common problem in testing multiple combinations of cluster locations and sizes), (d) its ability to identify the spatial locations where the null hypothesis is rejected, and (e) its ability to detect multiple clusters (22, 23).

Using Monte Carlo techniques, SaTScan assigned relative risk probabilities to defined block groups. Race and the standard age (18 years) groups were used as covariates in this analysis. Under the null hypothesis, the incidence of breast cancer follows a Poisson distribution, and the probability of a case being diagnosed in a particular location is proportional to the covariate-adjusted population in the location.

For hypothesis testing, the SaTScan program generated 999 random replications of the data set under the null hypothesis. The test statistic was calculated for each random replication as well as for the real data set. When the latter was among the 5% highest, the test was significant at the 0.05 level (24).

ArcGis version 9.0 was the geographic information system used for this analysis to view, analyze, and relate data from a spatial (geographic) perspective.

**Statistical Analysis.** All area-based measures derived from either the U.S. census data or the FCDS were merged with the master block group data file based on the block group key ( $n = 9,112$ ). Additionally, the results of the SaTScan analysis were merged with the master block group data file, identifying each block group as having a higher-than-expected incidence of late stage breast cancer or an expected incidence of late stage breast cancer.

All analysis was done on the master block group file using SPSS version 11.0.1 to analyze the degree to which SES was associated with breast cancer stage in Florida. Multivariate logistic regression was used to assess potential confounding variables (such as insurance, urban/rural location, and mammography usage) across groups.

## Results

The incident late stage breast cancer data set represented women from each county in Florida, with an age distribution of 24 through 85+ years old. The racial distribution of the cases was similar to the racial distribution of the state, which is 87.6% White, 10.7% Black, and 1.7% others. Once the data were aggregated by block group, women from 6,211 of the 9,112 block groups in Florida were represented in the analysis.

The study population was composed of 16,626 White women and 2,057 Black women. At the time of diagnosis, insurance status varied by racial group. Black women were approximately thrice more likely to be uninsured than White women (12.3% versus 4.7%) and 3.5 times more likely to receive

Medicaid (11.2% versus 3.3%). Based on block group area-based socioeconomic measures, the median household income for Blacks was approximately \$7,000 per annum less than the White population. Nearly one fourth of the Black women diagnosed with late stage breast cancer lived in areas of severe or near poverty as opposed to 7.5% of White women. The overall prevalence of mammography usage in Florida between Whites and Blacks is not statistically significant. The prevalence for White women is 79.7 [95% confidence interval (95% CI), 78.3-81.2] and 76.7 (95% CI, 71.1-82.3) in Black women.

SaTScan identified 767 block groups in which the incidence of late stage breast cancer cases was higher than expected. By default, all other block groups in the analysis file ( $n = 5,444$ ) were designated as having an expected incidence. The map of Florida (Fig. 2A) presents the areas that SaTScan identified as having a higher-than-expected incidence of late stage breast cancer.

To understand Florida's population and the nature of the different areas within Florida, SPSS and the spatial analytic properties of ArcGIS were used to compare and contrast the area-based characteristics of the different neighborhoods.

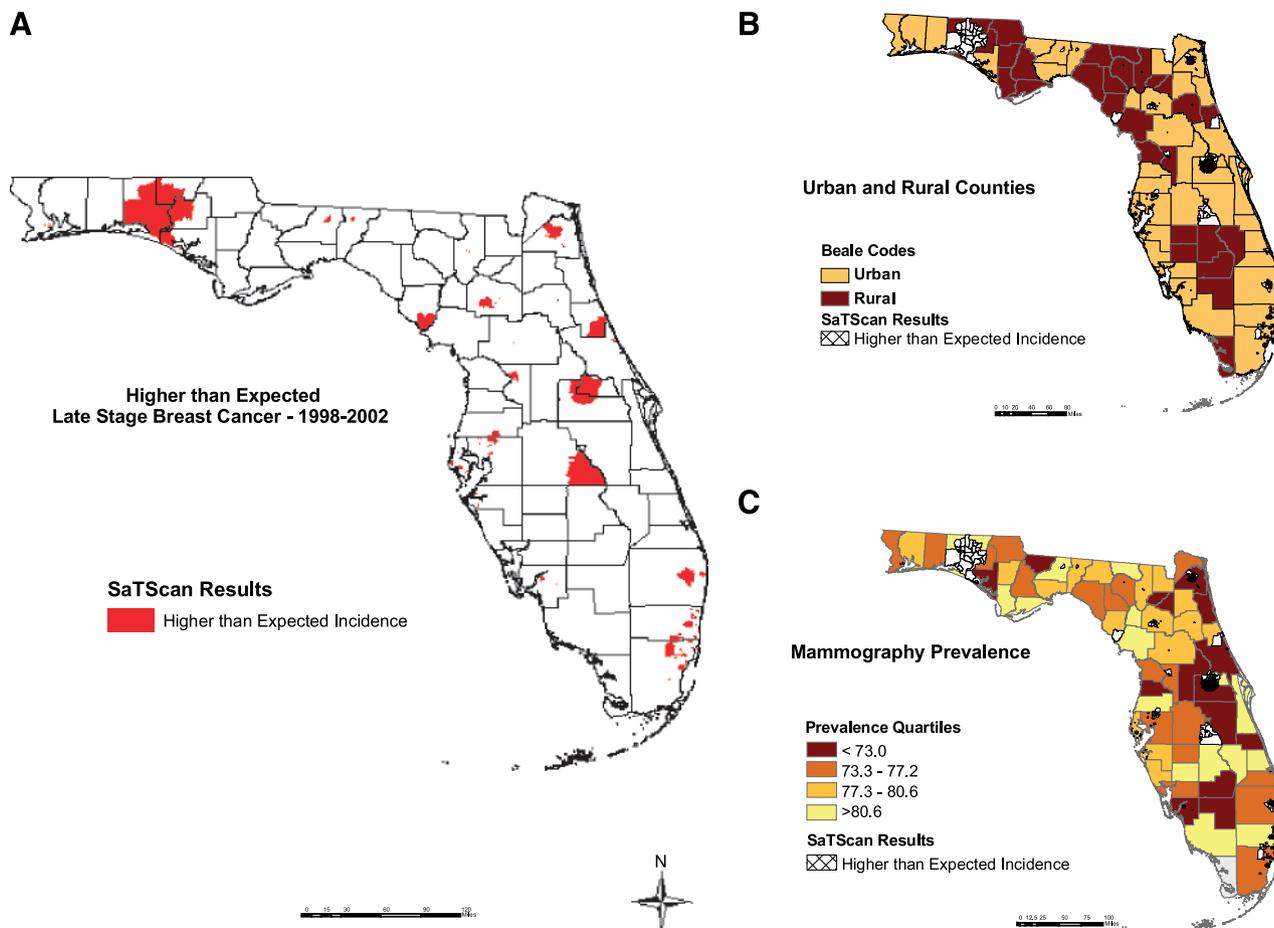
As seen in Fig. 2B, the distribution of urban and rural areas within the areas identified by SaTScan as having higher-than-expected incidence or expected incidence was striking. Over 95% of the areas identified as having higher-than-expected rates were located in urban counties ( $\chi^2 = 7.2$ , 1df,  $P = 0.007$ ). Although the areas identified by SaTScan as having a higher-than-expected incidence covered a relatively large geographic area, the population within these areas represented ~1.5 million (~9% of the total Florida population).

The logistic regression results (Table 1) present the odds ratios (OR) of SES on the areas with higher-than-expected

incidence versus expected incidence of late stage breast cancer. The table presents five models assessing the effect of SES on the diagnosis of late stage breast cancer, both as a simple regression and when area-based variables were sequentially added to the model as possible confounders. There were no significant interactions present in any of the models.

The effect of SES on the higher-than-expected incidence of late stage breast cancer was seen in women living in severely poor areas (OR = 2.6; 95% CI, 1.9-3.4) and women living in areas of near poverty (OR = 1.4; 95% CI, 1.0-2.2; model 1). Insurance and/or the urbanization of the area had little or no effect on this gradient (models 2-4). Whereas the magnitude of the effect is less for women living in near poverty, the implications were the same. Women living in near poverty were 50% more likely to be diagnosed with a late stage breast cancer than women residing in nonpoor areas.

The addition of mammography prevalence to the final model (model 5) had a significant effect on several variables in the model. The magnitude of the effect for both the severe and near-poverty categories increased. Areas of near poverty reached statistical significance (OR = 1.6; 95% CI, 1.1-2.6) and areas designated as severe poverty were thrice more likely to be an area of higher-than-expected incidence of late stage breast cancer (OR = 3.0; 95% CI, 2.2-4.0). Another striking result of this model was the gradient among the mammography prevalence quartiles. Areas with the lowest prevalence of mammography usage (<73.0) were 6.5 times more likely to be identified as having a higher-than-expected incidence of late stage breast cancer than areas with the highest mammography usage (>80.6; Fig. 2C). Likewise, areas in the second lowest quartile (73.0-77.2) were over 3.5 times more likely to be



**Figure 2.** Areas of higher-than-expected late stage breast cancer overlaid with urban/rural counties and mammography prevalence.

**Table 1. Logistic regression: ORs of living in areas of higher-than-expected incidence (as determined by SaTScan) of late stage breast cancer**

	OR	95% CI
<b>Model 1</b>		
SES		
Nonpoverty	1	—
Near poverty	1.5	1.0-2.2
Severe poverty	2.6	1.9-3.4
<b>Model 2</b>		
SES		
Nonpoverty	1	—
Near poverty	1.5	1.0-2.3
Severe poverty	2.6	1.9-3.5
Insurance		
Uninsured	1	—
Private	1.0	0.7-1.5
Medicare	1.0	0.7-1.8
Medicaid	0.7	0.5-1.1
<b>Model 3</b>		
SES		
Nonpoverty	1	—
Near poverty	1.4	0.9-2.2
Severe poverty	2.5	1.9-3.4
Urban/rural residence		
Rural	1	—
Urban	1.6	1.1-2.3
<b>Model 4</b>		
SES		
Nonpoverty	1	—
Near poverty	1.5	0.9-2.2
Severe poverty	2.6	1.9-3.6
Insurance		
Uninsured	1	—
Private	1.0	0.7-1.6
Medicare	0.7	0.5-1.1
Medicaid	1.1	0.7-1.9
Urban/rural residence		
Rural	1	—
Urban	1.5	1.1-2.2
<b>Model 5</b>		
SES		
Nonpoverty	1	—
Near poverty	1.6	1.0-2.6
Severe poverty	3.0	2.2-4.0
Insurance		
Uninsured	1	—
Private	0.9	0.6-1.4
Medicare	0.6	0.4-0.9
Medicaid	0.9	0.5-1.6
Urban/rural residence		
Rural	1	—
Urban	2.9	2.0-4.2
Mammography prevalence		
Quartiles (highest to lowest)		
4th, >80.6	1.0	—
3rd, 77.3-80.6	1.3	1.1-1.6
2nd, 73.0-77.2	3.6	3.0-4.5
1st, <73.0	6.5	5.1-8.3

identified as having a higher-than-expected incidence of late stage breast cancer (OR = 3.6; 95% CI, 3.0-4.5).

Unlike the previous four models, medicare was statistically significant in the final model (relative to uninsured), indicating that areas of higher-than-expected incidence of late stage breast cancer are 40% less likely to contain medicare recipients (OR = 0.6; 95% CI, 0.4-0.9).

## Discussion

The results of this research, indicating that residing in areas with high levels of poverty is significantly associated with late stage diagnosis of breast cancer, are noteworthy and consistent with other studies (25-30). These findings remained after attempting to control for important risk factors, such as insurance and mammography prevalence. In fact, the magni-

tude of the main effect was strengthened with the addition of mammography prevalence.

Insurance status was not significant in any of the models until the inclusion of mammography prevalence. With the inclusion of mammography prevalence estimates, medicare recipients were 40% less likely to live in areas of higher-than-expected incidence of late stage breast cancer. This may be due to the loss of specificity when the individual insurance variable was converted to an area-based measure.

The areas of higher-than-expected incidence of late stage breast cancer were almost seven times more likely to be the same areas with the lowest mammography usage. These findings are consistent with the well-documented benefits of mammography for the early detection of breast cancer. Despite this high association, it is also important to note that the area-based measure of poverty used in the present analysis remained statistically significant in our multivariate models, suggesting that mammography usage alone does not explain all of the variability in late stage breast cancer incidence rates in Florida.

The results of this study indicate that SES plays an independent role in the late stage diagnosis of breast cancer beyond the effects on mammography access. There are poorly understood (and sometimes controversial) underlying factors for increased late stage breast cancer incidence among women of lower SES that include access and or usage of health care (including cultural and social beliefs which may influence mammography usage; refs. 31, 32).

The "upstream" factors defining an individual's socioeconomic status are complex and multidimensional. Women of lower SES may be more likely to ignore symptoms for a variety of economic, social, or cultural reasons. These women tend to work in nonsalaried positions, therefore taking time off for a doctor's appointment more than likely results in a "no work-no pay" situation. Additionally, women of lower SES tend to be lesser educated and may not understand the elements of disease progression, and, as a result, they do not respond as quickly when feeling a palpable mass.

SES factors alone are often not sufficient to explain the dramatic variation in breast cancer stage among certain women. However, SES variables, in conjunction with cultural beliefs and attitudes, have been shown to account for a large portion of the observed effect (28). There is a growing body of literature suggesting an association between cultural, religious, and other psycho/social factors related to breast cancer screening (28, 33, 34). A dissertation of these factors is beyond the scope of this paper. However, the methods used in this research can assist cancer control professionals in identifying geographically defined subpopulations to survey for the identification of possible cultural, religious, and other psycho/social factor barriers to screening which could eventually result in the development of culturally sensitive interventions. Moreover, this methodology can offer researchers a baseline from which to compare the effectiveness of the interventions.

Results from the present study suggest that "where" you live plays an important role in defining the risk of presenting with late stage breast cancer. It is essential that additional research be conducted to assess and quantify barriers to health care, including, but not limited to, access and usage of primary health care in communities with higher-than-expected late stage cancer cases. Additional research is urgently needed to leverage the strengths and resources present in all communities to lower the late stage breast cancer burden, especially in high-poverty communities.

**Limitations.** There were several study limitations that should be noted, starting with an acknowledgement that cross-sectional study designs cannot be used to establish causal associations between SES and delayed breast cancer diagnoses. Also, because this is an ecological study, we must acknowledge the possibility of ecological fallacy.

The area-based population and SES classifications were assigned based on data collected in 1999 for the 2000 U.S. census. The population data were extrapolated for the 5-year study period by multiplying the single 2000 census year population by 5. To minimize the potential bias from this methodology, the study period was selected to include 2 years before and 2 years after the census year.

Estimates of mammography usage were available at the county, but not the block group level, from the Florida Behavioral Risk Factor Surveillance System. Use of county level data to estimate area level prevalence rates almost certainly resulted in some loss of precision given the likelihood of intracounty variability in these rates. Additionally, the Behavioral Risk Factor Surveillance System data may be overestimating recent mammography usage, particularly among low-income women (35).

A potential limitation could be attributed to the accuracy and completeness of the geocoded address data. Approximately, 94% of the cases were geocoded based on the complete address and street number. Only 5.3% of the cases were geocoded at only the zip code level (the centroid of the zip code was used to assign the geocoded variables). There were 309 cases excluded from the study because they were geocoded only to the county level, not at the block group level. There was nothing to indicate that any of the SES strata were disproportionately affected by either of these issues. If such a bias was to exist, it would be expected to effect each stratum equally.

Another potential limitation could have been the use of a single variable indicator versus composite indicators for the assignment of SES. However, based on data from Public Health Disparities Geocoding Project, Krieger et al. found that several of the single-variable measures, especially those intended to measure poverty, detected the same magnitude of socioeconomic inequality in health as the composite measures (36).

Finally, future analysis of these data should use multilevel modeling techniques that will take into account predictor and outcome variables that can occur at multiple levels of aggregation (i.e., the individual, block group, county, and state).

There are also several study strengths, including the use of spatial statistical technology to identify areas of interest below the county level and to incorporate race (in addition to age) as a covariate. Additionally, the inclusion of area-based socioeconomic measures allowed for the enhancement of cancer surveillance data, thus allowing for the inclusion of multifaceted indicators that are not routinely captured in surveillance systems.

## Summary

When analyzing overall breast cancer incidence, women of higher SES have a higher overall incidence of breast cancer. However, as was found in this study, when the analyses were stratified by the stage of the disease, there is a clear and significant gradient between late stage breast cancer and residence in high-poverty areas. Research has shown that usage of mammography screening in economically disadvantaged areas is less than wealthier areas for a variety of practical and economic reasons (37, 38). Therefore, additional research must be done to assess and understand the barriers to screening and/or the follow-up on positive screening examinations in these populations.

Health care access and usage may not be related solely to income and/or education. Social and economic factors, as well as cultural/spiritual beliefs, may be barriers to health care access and/or usage. To obtain this information, researchers must use emerging spatial methodology and survey techni-

ques that will enable them to augment the current surveillance systems. Additionally, the survey component would enable the research community to "verify" the area-based socioeconomic data.

There are many important reasons to continue enhancement of existing surveillance systems with area-based socioeconomic measures. The methodology used in this study not only gives public health practitioners the ability to identify communities that may benefit from targeted interventions to reduce their burden of delayed breast cancer diagnosis, but it is also a powerful tool for furthering our understanding of the role that SES indicators may play in complex disease processes. Additionally, this methodology has the ability to use modeling techniques to prospectively identify communities that have the potential to become areas of higher-than-expected incidence of cancers. Ultimately, this information can be used to develop effective and culturally sensitive interventions designed to reduce the cancer burden in all communities.

## References

- Krieger N, Williams DR, Moss NE. Measuring social class in public health research: concepts, methodologies, and guidelines. *Annu Rev of Public Health* 1997;18:341-78.
- Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev* 1988;10:87-121.
- Lynch JW, Kaplan GA, Cohen RD, Tuomilehto J, Salonen JT. Do cardiovascular risk factors explain the relation between socioeconomic status, risk of all-cause mortality, cardiovascular mortality, and acute myocardial infarction? *American Journal of Epidemiology* 1996;144:934-42.
- Breast Cancer Facts and Figures. Available from: <http://www.cancer.org/downloads/STT/CAFF2005BrF.pdf>.
- Robert SA, Strombom I, Trentham-Dietz A, et al. Socioeconomic risk factors for breast cancer: distinguishing individual- and community-level effects. *Epidemiology* 2004;15:442-50.
- Yabroff KR, Gordis L. Does stage at diagnosis influence the observed relationship between socioeconomic status and breast cancer incidence, case-fatality, and mortality? *Soc Sci Med* 2003;57:2265-79.
- Robsahm TE, Trelli S. Weak associations between sociodemographic factors and breast cancer: possible effects of early detection. *Eur J Cancer Prev* 2005;14:7-12.
- Smigal C, Jemal A, Ward E, et al. Trends in breast cancer by race and ethnicity: update 2006. *CA Cancer J Clin* 2006;56:168-83.
- Baquet CR, Commiskey P. Socioeconomic factors and breast carcinoma in multicultural women. *Cancer* 2000;88:1256-64.
- Evans T, Whitehead M, Diderichsen F, et al. *Challenging inequalities in health: from ethics to action*. New York: Oxford University Press; 2001.
- Kreiger N, Williams D, Moss N. Measuring social class in US public health research: concepts, methodologies and guidelines. *Annu Rev Public Health* 1997;18:341-78.
- Kreiger N. Overcoming the absence of socioeconomic data in medical records; validation and application of a census-based methodology. *Am J Public Health* 1992;82:703-10.
- Kreiger N, Chen JT, Waterman PD, Soobader MJ, et al. Choosing area-based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: the public health disparities geocoding project (US). *J Epidemiol Community Health* 2003;57:186-99.
- Kreiger N, Chen JT, Waterman PD, et al. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter? *Am J Epidemiol* 2002;156:471-82.
- Richardson LC, Tian L, Voti L, et al. The roles of teaching hospitals, insurance status, and race/ethnicity in receipt of adjuvant therapy for regional-stage breast cancer in Florida. *Am J Public Health* 2006;96:160-6.
- Voti L, Richardson LC, Reis IM, Fleming LE, Mackinnon J, Coebergh JW. Treatment of local breast carcinoma in Florida: the role of the distance to radiation therapy facilities. *Cancer* 2006;106:201-7.
- Poverty Thresholds Available from: <http://www.census.gov/hhes/www/poverty/threshld/thresh00.html>.
- Krieger N, Williams DR. Changing to the 2000 standard million: are declining racial/ethnic and socioeconomic inequalities in health real progress or statistical illusion? *Am J Public Health* 2001;91:1209-13.
- Beale codes. Available from: <http://www.ers.usda.gov/data/RuralUrban-ContinuumCodes/>.
- Kuldorff M. Information Management Services, Inc. SaTScan v3.1: software for the spatial and space-time scan statistics, 2002. Available from: <http://www.SaTScan.org>.
- Kuldorff M. A spatial scan statistic. *Communication in Statistics Theory and Methods* 1997;26:1481-96.
- Sheehan TJ, Gershman ST, MacDougall LA, et al. Geographic assessment of

- breast cancer screening in towns zip codes and census tracts. *J Pub Health Manag Prac* 2000;6:48–57.
23. Sheehan TJ, DeChello LM, Kulldorff M, Gregorio DI, Gershman S, Mroszczyk M. The geographic distribution of cancer incidence in Massachusetts 1988 to 1997, adjusted for covariates. *Int J Health Geogr* 2004;3:17.
  24. Kuldorff M, Information Management Services, Inc. SaTScan v3.1: Software for the spatial and space-time scan statistics. <http://www.SaTScan.org>.
  25. Farley TA, Flannery JT. Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications. *Am J Public Health* 1989;79:1508–12.
  26. Richardson JL, Langholz B, Bernstein L, Burciaga C, Danley K, Ross RK. Stage and delay in breast cancer diagnosis by race, socioeconomic status, age and year. *Br J Cancer* 1992;65:922–6.
  27. Mandelblatt J, Andrews H, Kao R, Wallace R, Kerner J. Impact of access and social context on breast cancer stage at diagnosis. *J Health Care Poor Underserved* 1995;6:342–51.
  28. Lannin DR, Mathews HF, Mitchell J, Swanson MS, Swanson FH, Edwards MS. Influence of socioeconomic and cultural factors on racial differences in late-stage presentation of breast cancer. *JAMA* 1998;279:1801–7.
  29. Gregorio DI, Walsh SJ, Tate JP. Diminished socioeconomic and racial disparity in the detection of early-stage breast cancer, Connecticut, 1986–1995. *Ethn Dis* 1999;9:96–402.
  30. Selvin S, Merrill DW, Erdmann C, White M, Ragland K. Breast cancer detection: maps of 2 San Francisco Bay area counties. *Am J Public Health* 1998;88:1186–92.
  31. Bailey EJ, Erwin DO, Belin P. Using cultural beliefs and patterns to improve mammography utilization among African-American women: the Witness Project. *J Natl Med Assoc* 2000;92:136–42.
  32. Russell KM, Perkins SM, Zollinger TW, Champion VL. Sociocultural context of mammography screening use. *Oncol Nurs Forum* 2006;33:105–12.
  33. Baquet CR, Commiskey P. Socioeconomic factors and breast carcinoma in multicultural women. *Cancer* 2000;88:1256–64.
  34. Maloney N, Koch M, Erb D, et al. Impact of race on breast cancer in lower socioeconomic status women. *Breast J* 2006;12:58–62.
  35. Peek ME, Han JH. Disparities in screening mammography. Current status, interventions and implications. *J Gen Int Med* 2004;19:184–94.
  36. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Race/ethnicity, gender and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures—the public health disparities geocoding project. *Am J Public Health* 2003;93:1655–71.
  37. Phelan JC, Link BG, Diez-Roux A, Kawachi I, Levin B. Fundamental causes of social inequalities in mortality: a test of the theory. *J Health Soc Behav* 2004;45:265–85.
  38. Link BG, Northridge ME, Phelan JC, Ganz ML. Social epidemiology and the fundamental cause concept: on the structuring of effective cancer screens by socioeconomic status. *Milbank Q* 1998;76:375–402, 304–5.