



Published in final edited form as:

Health Place. 2011 January ; 17(1): 327–334. doi:10.1016/j.healthplace.2010.11.007.

Geographic Disparities in Late-Stage Breast Cancer Diagnosis in California

Tzy-Mey Kuo, PhD^a, Lee R. Mobley, PhD^a, and Luc Anselin, PhD^b

Lee R. Mobley: lmobley@rti.org; Luc Anselin: luc.anselin@asu.edu

^a RTI International, 3040 Cornwallis Road, Research Triangle Park, NC, 27709-2194

^b Arizona State University, Tempe, AZ, USA

Abstract

Using cancer registry data for the population of California women aged 67+ with breast cancers, we estimated random intercept logistic models to examine how two socio-ecological predictors (residential isolation and poverty) were associated with probability of late-stage diagnosis for breast cancer. Using the multilevel modeling results, we calculated fully-adjusted predicted probabilities associated with women in each Medical Service Study Area (MSSA) in California and classified the areas into two distinct groups: MSSAs with predicted rates below the 25th percentile (presumably the better outcome areas) and MSSAs with predicted rates above the 75th percentile (presumably the worse outcome areas) for two minority groups. Some areas had better outcomes for one group but worse outcomes for the other, suggesting that interventions to improve outcomes need different strategies for different groups in the same areas. Using information from geographic risk factors and multilevel modeling, this study informs interventions designed to reduce disparities in breast cancer outcomes.

Introduction

Invasive breast cancer (BC) is one of the predominant diseases in older women. Compared with women younger than 65 years of age, older women are 4 times more likely to be diagnosed with invasive BC and about 8 times more likely to die from the disease (Horner et al., 2009). In addition, between 2000 and 2007, more than 30% of the invasive BCs among older women were diagnosed after the cancer had spread to regional lymph nodes, beyond the primary site (i.e., regional stage), or had already metastasized (i.e., distant stage) (SEER, 2010). Breast cancer diagnosed at a later stage has significantly lower survival rate than that diagnosed at an earlier stage. The survival rate is only 23.3% for the distant stage and 83.5% for the regional stage, whereas the survival rate is 98.3% for the localized stage (Horner et al., 2009), which may contribute to the large mortality burden for older women with invasive BC. Because the population size and life expectancy of older women continues to increase, (Administration on Aging, 2006; Hetzel and Smith, 2001) it is likely that the morbidity and mortality burden for older women will continue to increase. Thus, to improve BC outcomes for older women it is urgent that we develop strategies targeted to certain higher-risk locations and tailored to resonate with their higher-risk population subgroups.

Corresponding author: Tzy-Mey Kuo, mkuo@rti.org, phone: (919) 541-6622, fax: (919) 541-7384.

Disclosure: No competing financial interests exist

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Existing studies have identified several personal risk factors and contextual factors related to late-stage BC diagnosis, which may provide useful information for intervention development. Personal factors that are associated with a higher probability of late-stage BC diagnosis include minority race/ethnicity, being unmarried, having fewer routine health care visits, having a greater number of comorbidities, being older, lower utilization of mammography screening, and being insured under Medicare fee-for-service (FFS), as compared with Medicare managed care plans (Badgwell et al., 2008; Galit et al., 2007; Keating et al., 2004; Kirsner et al., 2005; Lee-Feldstein et al., 2001; McCarthy et al., 1998; Randolph et al., 2002; Riley et al., 1999; Riley et al., 2008). Contextual factors that relate to later stage of BC diagnosis are lower area-household income and smaller proportion of residents with at least a high school degree (Keating et al., 2004; McCarthy et al., 1998). In addition, a recent study, conducted by Haas et al. (2008) found that residential isolation of racial or ethnic groups in combination with area-level income mediated the likelihood of late-stage BC diagnosis among older women across SEER Registry populations. They grouped the areas into 4 types based on an isolation index of residential segregation and area poverty. To define the areas, they used a cutoff point approximately equal to the 75th percentile of the isolation index for the areas used in the study. They dichotomized income using a threshold of 200% of the 1990 federal poverty threshold for 1 person, based on the median per capita income for the census tract where the sample member resided. Specifically, they found that African American or Hispanic women who lived in low-isolation and low-income areas (below 75th percentile isolation and median income less than 200% of the 1990 federal poverty threshold, respectively) had a greater probability of late-stage BC diagnosis than white women.

Several recent studies have used a combination of traditional statistical analysis and mapping to show geographic areas demonstrated by statistical modeling to have higher risk of late-stage BC that are in need of intervention (Goovaerts, 2010; Gumpertz et al., 2006; MacKinnon et al., 2007; McElroy et al., 2006; Roche et al., 2002). For example, Gumpertz et al. (2006) presented maps showing the spatial distribution of several significant predictors from their multivariate regression results. They found that, in Los Angeles County, advanced BC diagnosis was more likely in areas that had a larger proportion of minority racial or ethnic groups or low median household income. They provided maps showing areas with low median household income, or higher rates of advanced-stage BC at diagnosis by racial and ethnic group. A study by MacKinnon et al. (2007) used spatial cluster analysis to identify areas with higher than expected incidence of late-stage diagnosis. They then used logistic regression on the binary area type (clustered, not clustered) to examine associations between binary area type and contextual factors, finding that areas with higher poverty and lower mammography usage were more likely to be located in the high-incidence clusters.

The combination of findings from multivariate analysis and mapping of geographic location associated with greater risk for worse outcome can provide rich and useful information regarding contextual factors and places associated with late-stage diagnosis of BC. In this paper, we follow the tradition of showing the spatial distribution of places with worse outcomes identified in multivariate models. The main contribution of our paper is to use a random-intercepts multilevel model to generate a list of places ranked by predicted outcomes, and demonstrate that places may be characterized as higher risk for some groups but not others.

Methods

Study Sample

In this study, we use the California Surveillance, Epidemiology, and End Results (SEER) Registry population, restricted to include only women who had primary BC diagnosis at age

67 years or older during the years 2000 through 2005. For women who had multiple BC diagnoses, we selected only those whose first diagnosis occurred at age 67 or older. We also restricted the sample to those enrolled in traditional FFS Medicare, because the analysis controlled for mammography screening history in the 2 years prior to each woman's BC diagnosis, determined from Medicare claims files and available only for persons with traditional FFS Medicare coverage. We restricted the sample to those with traditional FFS Medicare in the month of and 2 years prior to the month of diagnosis, to ensure we had complete Medicare claims histories for the study sample. Furthermore, because we wanted to focus on factors related to BC as the first cancer diagnosis, we excluded women who had other types of cancer diagnosis prior to the BC diagnosis.

Data Sources

The SEER Registry provides rich information for the population with cancer diagnoses, including month and year of cancer diagnosis, cancer site, histological type, cancer stage information, ZIP code of residence, and patient demographics. These cases were linked to Medicare claims in the SEER-Medicare database used for this study, provided by the National Cancer Institute. We used residential ZIP code to merge in contextual variables from the public-use RTI Spatial Impact Factor Database (RTI, 2009). We used the contextual factors defined at the Medical Service Study Area (MSSA), defined by the state of California to break up counties into urban, rural, and frontier areas for the purpose of allocating scarce health care resources (Christman, 2004). MSSAs are much smaller than counties and nest inside them (Mobley et al., 2008).

Measures

Person-level variables—Stage of diagnosis was defined by the SEER summary staging system (i.e., in situ, localized, regional, distant, or unstaged). We included only women with invasive cancer diagnoses, grouping the regional and distant stages as “late stage” and the localized stage as “early stage.” Women with unstaged diagnosis were dropped from the analysis (N=1,261), leaving 33,838 women in the final analytic file. The majority of women with unstaged diagnosis were white (88.7%). The proportion of unstaged diagnosis was similar across racial and ethnic groups, ranging from 1.6% in Hispanics, 2.5% in Asians, 3.1% in whites, to 3.7% in African Americans.

We created an indicator variable for any mammography use in the 2 years prior to the date of first BC diagnosis, and included both screening and diagnostic types of mammography. To avoid including in the screening indicator any diagnostic mammography associated with the BC discovery, we excluded from the screening indicator claims for any mammography use that occurred 3 months or less from the month of BC diagnosis.

Demographic information on subjects included age, race/ethnicity, marital status, original reason for Medicare entitlement (which identifies persons with lifelong disabilities), and an indicator of whether the state provided assistance to a low-income woman in purchasing Part B insurance or covering copays and deductibles (i.e., a state buy-in recipient)(MEDPAC, 2004). Marital status and state buy-in recipient indicator were defined in the year of BC diagnosis. The original reason for Medicare entitlement variable was included to help control for comorbidities, indicating whether a woman had end-stage renal disease (ESRD) or disability at time of entitlement.

Geographic Risk Factors—One of the key contextual variables is the residential isolation measure defined by Massey and Denton (1988). Their five dimensions of residential segregation have been used extensively in the health outcomes literature (Kramer and Hogue, 2009). The five dimensions are evenness/dissimilarity, exposure/isolation,

concentration, centralization, and spatial clustering. **Evenness/Dissimilarity** reflects the degree to which each neighborhood within a larger area contains the same proportion of minorities and whites as the larger area overall; it reflects the uniformity of proportions of minorities across geography and can handle more than two groups in the calculation. **Exposure/Isolation** reflects the degree to which minority populations are exposed to other minorities rather than to whites, and a complementary measure reflects the opposite construct, the degree of interaction among whites and minorities. The exposure measures can handle only two groups and are often calculated separately for each race or ethnicity relative to whites. **Concentration** refers to the relative amount of physical space occupied by minority groups in areas, and reflects differences in population densities experienced by people of different races. **Centralization** captures the degree to which a group is spatially located near the center of an urban area. **Spatial Clustering** reflects the extent to which minorities are clustered into enclaves within larger areas, which may be a less useful construct outside urban areas where housing projects are infrequent, which is a concern here because our analysis sample spans the urban-rural continuum.

We use the isolation index in this paper because we are interested in capturing effects associated with social cohesion or support. We expect that social support is reinforced by the degree of contact with members of ones' own race or ethnicity in smaller areas, such as MSSAs (which are sub-components of counties). As shown in Mobley et al (2008) in California, multivariate models using the MSSA, primary care service area (PCSA), or ZIP code tabulation area (ZCTA) (all much smaller than counties) produced consistent associations between area residential isolation and individual mammography utilization, while the county measure of residential isolation produced associations inconsistent with (opposite in sign) these smaller areas. They argued that this reflects differences in interpretation of residential isolation indices when a person's neighborhood is defined as a small local area, reflecting social support, versus when areas are large, reflecting geo-political units and degree of political influence held by minorities.

This notion that findings differ because measures reflect different 'levels of influence' has been reinforced in the discussion by Kramer and Hogue (2009). Only 2 of the 39 studies reviewed by Kramer and Hogue (2009) used a spatial clustering measure, whereas 11 used an isolation index and 12 used a dissimilarity index. Several studies posit that isolation (among African Americans) reflects an adverse environment that can be modified positively by a high degree of clustering into enclaves which enhances political empowerment (Bell et al, 2006; Laveist, 1992; 1993). Bell et al (2006) find a negative effect of isolation and, conditional on that, a positive effect from clustering. However, these studies focused on urban areas, where the clustering index is more useful. Taking all things into consideration, and considering our desire to make our findings more comparable with recent literature while reflecting the exposure dimension we believe is important, specifically as regards reinforcing tailored communications aimed at minority populations by cancer control groups and activists, we chose to use the isolation index.

The residential isolation measure used here is defined at the MSSA neighborhood level, with values ranging from 0 to 1 and where higher values representing greater isolation or segregation of minorities from whites (Massey and Denton, 1988). This index is calculated separately for African Americans, Asians, and Hispanics relative to whites. Because residential isolation may be confounded with area-level poverty, following Haas et al. (2008) we created a four-level area type indicator for each race or ethnicity. The four area types are characterized by: high isolation and high poverty (HI-HP), high isolation and low poverty (HI-LP), low isolation and high poverty (LI-HP), and low isolation and low poverty (LI-LP). Isolation for each group (African American, Hispanic, and Asian) was dichotomized as "high isolation" if the area-level isolation measure was above the 75th

percentile among MSSAs in the state; otherwise, the areas were classified as “low isolation.” We chose the 75th percentile to be consistent with recent comparable work by Haas et al. (2008). We used percent of the population aged 65+ living below the poverty line as the poverty measure, and areas were classified as “high poverty” if the percent poverty was above the 75th percentile among MSSAs in the state; otherwise, the areas were classified as “low poverty.” It is worth noting that while above 75th percentile in poverty or isolation qualifies as “high,” one could argue that simply being below the cutoff point does not make it “low.” We chose these terms to be comparable to previous literature and to make interpretations easier.

We included several area-level contextual variables that were found to be significant predictors of mammography utilization in 2002–2003 (Mobley et al., 2008), thus could also be associated with the stage at diagnosis. This list of variables was derived from a comprehensive conceptual model of the various levels of influence for socio-ecological predictors. These variables include neighborhood measures defined at the MSSA unit: proportion of the area that is rural, proportion of the workforce that commute longer than 60 minutes to work (each way), mammography facility density (number of facilities per 1,000 women aged 40 or older), proportion of population that are foreign born and entered the United States from 1995 to 2000, and number of primary care physicians, obstetrician, or gynecologists per 1,000 population aged 65 or older. The county level of geography, reflecting broader markets influenced by Medicare program resources, is used for defining number of oncologists per 1,000 population aged 65 or older and Medicare managed care plan penetration.

Statistical Analysis—We modeled the probability of late-stage breast cancer diagnosis using patients diagnosed at an early-stage as the reference group in multilevel logistic regression. To account for possible spatial heterogeneity, we allowed the intercept of each MSSA to be different, using a random intercepts model. Other methods for dealing with potential spatial heterogeneity exist. Brundson et al (1999) described three methods including random intercepts, random intercepts and random slopes, or geographically weighted regression. In their study, the three methods produced different results for the same data set, and they concluded that there is no way to determine which approach is the ‘best’. We argue that ex-ante conceptualization of the ecological system is necessary before choosing an appropriate empirical modeling approach.

Our conceptual model describes spatial interaction among people and characteristics of their contextual environments along the pathways to healthcare utilization, which partly determines health outcomes (Mobley et al, 2008). In the conceptual model, each of the states in the US represents a unique healthcare environment governed by local and regional politics, social systems, and baseline conditions. Within the state there are market-level forces that determine supply factors, and community or neighborhood-level forces that determine social factors, and individuals with predisposing, enabling, and need characteristics which interact with the forces in the broader system. Such a comprehensive conceptualization includes all factors thought to be important contributors to the observed outcomes, which are measured at the person-level.

This conceptual model discounts the notion that the ecological system itself varies from place to place within the state; it maintains that this system does vary across the states. Thus we do not believe that a geographically-weighted regression (GWR) or a random coefficients (intercepts and slopes) model is appropriate. However, to the extent that we have omitted important contextual variables from the model, it may be necessary to use a random intercepts formulation for the multilevel model (with person and area-level covariates). In this regard, our model is very similar to the one used by Gumpertz et al

(2006) to model advanced-stage breast cancer incidence in Los Angeles county. As described in Gumpertz et al (2006) and in Oakes (2004), a small variance estimate for the area-level random effect indicates that the contextual factors included in the model do a good job accounting for spatial heterogeneity in the explanatory factors.

A main advantage of the random-intercepts models is their utility in predicting area-level rates for the outcome of interest (prevalence of late-stage diagnosis) that account for differences in sample sizes and selection (migration) of people into areas. As noted by Gelman and Hill (2007, p. 271), in situations where the variance estimate for area-level effects is small, a multilevel approach is most important because it allows intercept estimates to vary by group yet estimates them precisely, especially for groups with small sample size. More specifically, when there are no omitted area-level effects and if people were randomly assigned to areas, any variation in health outcomes between areas could be attributed to the areas themselves (Oakes, 2004, p. 1934). However, people are not randomly assigned to areas; they self-select their locations so it is important to include variables such as age, marital status, comorbidity, and socioeconomic status. When these variables determine selection behavior, then by including them in the multilevel model the predicted person-level outcomes (used to construct area-level rates) are adjusted for differences in neighborhood composition due to selection. Thus the observed differences in health outcomes between areas cannot be separated from the selection or biological information of people residing there (Oakes, 2004: pp 1938–1939). This renders the predicted area-level estimates of rates of late-stage diagnosis as particularly meaningful for site selection of locations for interventions, because they reflect both the compositional characteristics of the resident populations and the contextual factors characterizing conditions in the area.

In this study, we estimate a random intercepts model with the following functional form:

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_i + \gamma_{01}W_j + \gamma_{02}C_j + \gamma_{20}R_i + \gamma_{21}R_i * C_j + \mu_{0j} + e_{ij}$$

Where Y_{ij} is log of the late-stage probability divided by early-stage probability for person i in MSSA j , X_i represents all personal level factors except race or ethnicity, R_i is a binary indicator for race or ethnicity (African American, Hispanic, or Asian, relative to white), W_j represents the contextual factors characterizing MSSA areas j , C_j is the indicator for the four different isolation-poverty area types, e_{ij} is the residual of the outcome variable Y_{ij} , and μ_{0j} is the unique effect of MSSA area j on the outcome and is a random variable. The multilevel model was estimated separately for whites and African Americans, for whites and Hispanics, and for whites and Asians, using white as the comparison group in each model. The random intercept models were estimated using GLIMMIX procedure in SAS version 9.1.3 (SAS Institute, Cary, NC).

Results

Sample statistics are presented in Table 1. African Americans and Hispanics were more likely to have a late-stage BC diagnosis than whites, and Asians were less likely than whites to have a late-stage BC diagnosis. Asians had higher mammography utilization and Hispanics had lower utilization in the 2 years prior to BC diagnosis as compared to whites and African Americans. MSSA-level characteristics are presented in Table 2. The median residential isolation index was higher for Hispanics than for African Americans or Asians.

Using the random intercepts multilevel models, we found that the variance of the random intercepts (μ_{0j}) was near zero (0.0025, 0.0016, and 0.0022 for the African American/white, Hispanic/white, and Asian/white models, respectively), indicating that spatial heterogeneity

modeled well by the personal and contextual factors. Table 3 presents the coefficient estimates from these models. Across all models, women aged 70 to 84 were significantly less likely to have invasive BC diagnosed at a late stage than women in younger (age 67–69) or older (age 85+) groups. Married women were less likely and women with state buy-in status were more likely to have a late-stage invasive BC diagnosis. Although the state buy-in variable is not a perfect measure of personal income, because not all poor women apply for this low-income subsidy, for those women who do receive it we are confident that personal income levels are low. Low personal income was associated with much higher probability of late-stage diagnosis across all three models. Controlling in this way for individual-level low income, the area-level poverty variable was more likely to represent an area-level effect.

The African American (versus white) and Hispanic (versus white) models showed significant race-by-isolation–poverty interaction, whereas the Asian (versus white) model did not (Table 3). The overall F-test for the race/ethnicity by area type interaction were: $F = 4.65$; $p=0.0035$ (African American), $F = 2.74$; $p=0.0434$ (Hispanic) and $F = 0.21$; $p=0.8872$ (Asian, data not shown). Although the overall interaction effect of race by area type was significant in both African American and Hispanic models, the significant effects did not occur for the same area types (HI-HP, HI-LP, LI-HP, LI-LP). Thus to better demonstrate the significance of the race/ethnicity by area type interaction effects for these two models, we used the area type associated with the best outcome as the reference group in each model. Because the results from the multivariate analysis for the Asian group did not show a significant interaction effect between race and isolation–poverty, and because the predicted marginal probabilities for Asians were smaller than whites in all four classes of isolation–poverty (data not shown), we do not further examine results for this group’s model.

To facilitate interpreting the net numerical effect of a binary race variable, an isolation–poverty variable, and their interaction, we summarized the findings in terms of predicted marginal probabilities of late-stage outcomes for women in the four area types (Lane and Nelder, 1982), obtained through LSMEANS option of the SAS GLIMMIX procedure. These are presented in Figures 1 and 2. In Figure 1, the probability of late-stage BC among white women with invasive breast cancer was about 35%–39%, with those living in high-poverty areas having slightly higher probabilities. For African Americans living in the high poverty areas, the disparity between African American and white in the probability of late-stage breast cancer was nearly diminished in areas with high African American isolation; but in low isolation areas, there was about a 5% higher probability of late-stage diagnosis for African Americans versus whites (Figure 1). Thus higher isolation (from whites) seems to confer social support for African Americans in high-poverty communities. While disparities between African American and white outcomes were larger in the non-poor communities, even in these communities African Americans living in greater isolation had better outcomes than less isolated individuals. Thus the ‘best outcome area type’ for African Americans was found to be in areas with high African American isolation and high poverty (HI-HP), where the probability of late stage diagnosis was lowest and roughly commensurate with whites. The ‘worst outcome area type’ for African Americans had low African American isolation and low poverty (LI-LP)(Figure 1). The difference between the best and worst outcome area types for African Americans relative to whites was significant ($p = 0.042$ after Bonferroni correction)

As shown in Figure 2, we found that when Hispanic women lived in high-poverty areas and were not exposed to many other Hispanic women (LI-HP), the disparity in their outcomes relative to whites was the largest (i.e. the ‘worst’ outcome area types for Hispanics) – with a 47% probability of late stage diagnosis for Hispanics (versus 36% for whites). However, the disparity between Hispanics and whites in high-poverty areas with high Hispanic isolation was very small (HI-HP, 38% probability for Hispanics and 37% probability for whites).

These findings suggest that in high-poverty areas, the exposure of Hispanic women to other Hispanics helps ameliorate the deleterious effects of living in impoverished neighborhoods. In addition, for the low-poverty areas, Hispanic women living in greater isolation/more exposure among other Hispanics (i.e., HI-LP) had lowest probability (30%) of late-stage BC diagnosis (i.e., the best outcome areas) which was smaller than the probability of late-stage BC diagnosis for the Hispanic women living in low isolation-low poverty (LI-LP) areas (38%). The difference between the best and worst outcome area types for Hispanics relative to whites was marginally significant ($p = 0.102$ after Bonferroni correction). Thus irrespective of poverty level in areas, Hispanics in residential enclaves with other Hispanics received some sort of social benefits which improved their health outcomes.

Spatial Translation of Findings

The area type interaction estimates summarized in Figures 1 and 2 are population averages, most precise at the point of means, and as such they represent average tendencies or associations in the data. We are confident that areas with certain isolation-poverty characteristics are associated with better or worse outcomes for minorities, on average. However, the challenge is to find a way to determine the most at-risk communities for each minority from the regression results. These will not likely be all areas characterized by particular combinations of isolation-poverty because effects ‘on average’ can’t identify specific places.

To identify MSSA areas with higher risk of late-stage BC at diagnosis, we used all of the multilevel modeling results, including the random intercepts, to characterize and rank MSSA areas by risk of late stage diagnosis. This is the main advantage of estimating this type of model (Gelman and Hill, 2007), as discussed above. To do this, we used the random intercepts multilevel model to compute predicted probabilities of the outcome for each person, fully adjusted for all area and person-level variables. We then aggregated these to the MSSA level for each race and ethnicity. We then ranked the MSSAs from highest to lowest and selected MSSAs with less than 25th percentile of the predicted probabilities (‘low’ areas, presumably the better outcome areas) and MSSAs with greater than 75th percentile of the predicted probabilities (‘high’ areas, presumably the worse outcome areas) for each minority group.

We identified 65 low- and 64 high-areas for African Americans and 79 low- and 79 high-areas for Hispanics. From these areas, we found that the MSSAs with the highest and lowest predicted rates of late-stage diagnosis were qualitatively and geographically different for the two minority groups. Comparing the low- and high-areas for these two minority groups, we found about 20 areas where the probability of late-stage BC diagnosis was high for African Americans but low for Hispanics and vice versa. In addition, 4 areas had low probability of late-stage BC diagnosis for both minority groups and 5 areas had high probability for both minority groups, with about 12% difference in the probability of late-stage BC diagnosis between these high and low areas for both groups. The locations of these four types of identified areas are shown in Figure 3. These areas span the urban-rural continuum across the state of California, and can be identified exactly from the list of places used in the GIS. Using our methods to compare areas with different predicted outcomes between the minority groups can identify a subset of locations where outcomes for the groups are different, suggesting interesting areas for further study. We present the identified MSSAs in Figure 3, created using ArcGIS (ESRI, Redlands, CA, version 9).

Discussion

Our results from multilevel regression analysis are largely consistent with the literature for the personal level predictors. We found that unmarried women and state buy-in recipients

had higher proportion of late-stage BC diagnosis than their counterparts. Our finding that women aged 67–69 were more likely to be diagnosed with BC at a later stage than those 70 and older highlights the importance of BC screening for this particular age group. According to a report from the state of California, women aged 65–69 had the highest age specific incidence rate of invasive breast cancer between 1998 and 2000 (California Cancer Registry, 2004). This pattern is consistent across all racial or ethnic groups in California except for Asians, where the highest incidence rate is for women aged 60–64. Other recent evidence suggests that women 65 and older experienced the sharpest decline in mammography screening rate from 2000 to 2005 nationwide (Breen et al, 2007). Thus, strategies to promote BC screening for California women ageing-in to Medicare (age 65) seem particularly important.

Our findings are different from Haas et al. (2008) for the effect of isolation-poverty areas on the late-stage BC diagnosis for African Americans and Hispanics. Haas et al. (2008) found people living in LI-HP areas were associated with higher late-stage BC rates for both African Americans and Hispanics. We find no such consistency among African American and Hispanic groups in California. There are several possible explanations for the difference between their study and ours. First, our definition of late-stage BC diagnosis includes regional and distant tumors, whereas theirs included only distant tumors. However, we re-ran the regression analyses using their definition and found exactly the same worst outcome area types for African American and Hispanic as identified by our original definition. In addition, they pooled all 12 states with SEER Registries together, whereas we only examined California. The California ecological system may be different than an average across 12 states, which demonstrate the value of analyzing each state separately when statistical power permits. The set of predictors in their paper was less comprehensive than ours, so unmodeled spatial heterogeneity may have impacted the findings. Finally, they defined a mixture of counties and smaller municipal areas as the spatial units for their isolation measure, and income was defined at a person's census tract. We defined both these contextual variables at the California MSSA area level.

Our model suggests that for both African American and Hispanic minority groups, whether living in poor or non-poor communities, residential isolation among those of the same race/ethnicity reduces the probability of bad outcomes. These findings supports the concept that isolation (higher probability of contact with persons of own minority type than whites) can confer social support that improves health outcomes. This finding is consistent with recent literature which digs deeper into the potential sources of social benefits, positing that ethnic enclaves provide protective health benefits by centralizing neighborhood and social or cultural resources (Finch, Lim, Perez, & Do, 2007). Others have looked at assimilation of new immigrants into communities over time, finding differences in social forces among new immigrant and established communities (Keegan, et al. 2010). Findings of different enclave effects for new immigrants versus US-born persons may be especially salient in California, where large populations of both types reside. It is likely that the older SEER population that we studied were residents of the latter type (i.e., US-born) communities, whose residents may be better educated or more acculturated into health policy and opportunities under Medicare. On the other hand, the low isolation/high poverty neighborhood findings may reflect the downward progress made by U.S. born Mexicans, who are able to spatially assimilate into less isolated neighborhoods but are unable to do so economically.

In addition, we found that there may be a quite beneficial effect associated with African Americans living in greater residential isolation among other African Americans in the MSSAs with higher poverty (the HI-HP area type). This finding suggests that there have been successful intervention efforts resonating with lower-income African American women living in segregated communities in California. For Hispanics, the most beneficial effect was

found in the MSSAs with greater residential isolation for Hispanics and lower poverty (the HI-LP area type). As suggested by recent studies regarding assimilation of Hispanic immigrants, these enclaves are likely populated with better educated Hispanics (Keegan, et al. 2010).

Conversely, MSSAs with LI-LP for African Americans seem to be the most detrimental to African American outcomes, whereas the most detrimental communities for Hispanics are those with LI-HP. Low residential isolation is thus associated with worse outcomes while high residential isolation is associated with better outcomes, for both groups. This suggests that residential isolation (higher probability that minorities associate with minorities) confers a degree of social support in California (Kramer and Hogue, 2009). Or, it could reflect more successful interventions employed in minority enclaves than elsewhere, which is a fruitful topic for evaluation of interventions in comprehensive cancer control.

The isolation-poverty area types defined in this study used the cutoff point of 75th percentile for residential isolation and poverty defined at the MSSA level. We conducted a sensitivity analysis using the 65th, 75th, and 85th percentiles as the classification cutoff points for both variables, and found that the main findings for this study, summarized in Figures 1 and 2, are not sensitive to changing these thresholds. Irrespective of the threshold used, the relative rankings of predicted probability by poverty-isolation area type are robust.

This study has several limitations. First, the study sample is drawn from the population of California women aged 67 years or older with BC who are also insured by FFS Medicare. The results cannot be generalized to women with other forms of health insurance, younger women, or women living in other states. Second, an assumption with contextual factors is that women lived long enough in an MSSA to be impacted by factors that characterized this MSSA. On the other hand, especially in higher-poverty neighborhoods, high mobility rates may be the norm within some area types. Third, there are many dimensions governing the complexities inherent in modeling and understanding social determinants of disparities (Osypuk and Acevedo-Garcia, 2010). While it is beyond the scope of this paper to explore all possibilities, some data to explore these things are readily available (RTI, 2009) and may be useful in future research exploring interactions of persistent poverty, income disparity, and various segregation measures, contributing even more to what we know about the social causes of disease.

Conclusions

This paper demonstrates an approach that combined findings from multilevel modeling and mapping of geographic locations to provide information regarding contextual factors and places associated with higher or lower rates of late-stage diagnosis for BC. Specifically, we used the random intercepts form of multilevel logistic regression models to estimate the likelihood of late-stage of breast cancer diagnosis separately for each race or ethnicity relative to whites. The random intercepts estimator accounts for 'compositional' characteristics of people in MSSAs to yield estimates of MSSA-level rates of late stage diagnosis that are adjusted for sample composition as well as sample size. The predicted rates of late-stage diagnosis from the model are thus smoothed relative to the distribution of raw rates. The precision of the point estimates used in aggregate to form the area rates may be variable, however we regard the estimates as an improvement over the raw rates or other methods (Haomiao et al, 2004), which have extreme variability related to both sample selection and sample size. Using the estimates, we identified MSSA areas where the predicted rates of late-stage outcomes were in the upper and lower quadrants of the distribution. We focused on the subset of specific MSSAs, identified as 'high' (upper 75th percentile) or 'low' (lower 25th percentile), that were identified for both minority groups.

We found some of these areas had high predicted rates for one racial/ethnic group but low predicted rates for the other racial/ethnic group, indicating interesting geographic areas for further study. We also found areas with high or low predicted rates for both groups. Using GIS to identify these specific areas, along with their predicted outcomes rankings and underlying socio-ecological data, targeted interventions can be planned that are informed by these findings. Contrasting areas known to have different outcome patterns among the groups is a fruitful way to design focused intervention studies with enough variability across locations to provide meaningful or statistically significant insights from case studies using mixed methods research. Thus researchers and intervention planners using this approach can decide how to most effectively and efficiently allocate resources or tailor interventions to fit subpopulations with known social conditions, so as to ultimately reduce geographic variation of late-stage BC diagnosis.

Acknowledgments

This work was supported by a National Cancer Institute grant (1R01CA126858-01A1) and RTI International Professional Development Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of RTI International, the National Cancer Institute, or the National Institutes of Health.

References

- Administration on Aging. Federal interagency forum on aging-related statistics: older Americans update 2006: key indicators of well-being. United States Government Printing Office; Washington, DC: 2006.
- Badgwell BD, Giordano SH, Duan ZZ, Fang S, Bedrosian I, Kuerer HM, Singletary SE, Hunt KK, Hortobagyi GN, Babiera G. Mammography before diagnosis among women age 80 years and older with breast cancer. *Journal of Clinical Oncology*. 2008; 26:2482–2488. [PubMed: 18427152]
- Bell JF, Zimmerman FJ, Almgren GR, et al. Birth outcomes among urban African-American women: a multivariate analysis of the role of racial residential segregation. *Soc Sci Med*. 2006; 63(12):3030–3045. [PubMed: 16997438]
- Breen N, Cronin KA, Meissner HI, Taplin SH, Tangka FK, Tiro JA, McNeel TS. Reported drop in mammography: Is This Cause for Concern? *Cancer*. Jun; 2007 109 (12):2405–2409. [PubMed: 17503429]
- Brundson C, Aitkin M, Fotheringham S, Charlton M. A comparison of random coefficient modeling and geographically weighted regression for spatially non-stationarity regression problems. *Geographical & Environmental Modeling*. 1999; 3(1):47–62.
- California Cancer Registry, Cancer in California. 1998–2000 Section III Breast Cancer. accessed on February 24, 2010 <http://www.ccrca.org/Cancer05/index.html>
- CDC. MMWR Surveillance Summaries, Special Focus: Behavioral Risk Factor Surveillance —United States. 1993 Aug 27.42(SS-4) 1991.
- Christman, JS. Working paper. University of Southern California, School of Policy, Planning, and Development; 2004. Urban and rural designations: The impact on rural healthcare in California. <http://rtispatialdata.rti.org> rest of link: <http://rtispatialdata.rti.org/LinkClick.aspx?fileticket=zSU2j8hGb4c%3d&tabid=64&mid=391>
- Finch BK, Lim N, Perez W, Do P. Towards a population health model of segmented assimilation: The case of low birth-weight in Los Angeles. *Sociological Perspectives*. 2007; 50(3):445–468.
- Galit W, Green MS, Lital K. Routine screening mammography in women older than 74 years: A review of the available data. *The European Menopause Journal, Maturitas*. 2007; 57:109–119.
- Gelman, A.; Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press; 2007.
- Goovaerts P. Visualizing and testing the impact of place on late-stage breast cancer incidence: a non-parametric geostatistical approach. *Health & Place*. 2010; 16:321–330. [PubMed: 19959392]

- Gumpertz ML, Pickle LW, Miller BA, Bell BS. Geographic patterns of advanced breast cancer in Los Angeles: associations with biological and sociodemographic factors (United States). *Cancer Causes and Control*. 2006; 17:325–339. [PubMed: 16489540]
- Haas JS, Earle CC, Orav JE, Brawarsky P, Neville BA, Williams DR. Racial segregation and disparities in cancer stage for seniors. *Journal of General Internal Medicine*. 2008; 23:699–705. [PubMed: 18338215]
- Hetzel, L.; Smith, A. The 65 years and over population: 2000. Washington, DC: United States Census Bureau; 2001. census 2000 brief, C2KBR/01–10
- Haomiao J, Muennig P, Borawski E. Comparison of small-area analysis techniques for estimating county-level outcomes. *American Journal of Preventive Medicine*. 2004; 26(5):453–460. [PubMed: 15165663]
- Horner, MJ.; Ries, LAG.; Krapcho, M.; Neyman, N.; Aminou, R.; Howlader, N.; Altekruse, SF.; Feuer, EJ.; Huang, L.; Mariotto, A.; Miller, BA.; Lewis, DR.; Eisner, MP.; Stinchcomb, DG.; Edwards, BK. SEER cancer statistics review, 1975–2006. National Cancer Institute; Bethesda, MD: 2009. http://seer.cancer.gov/csr/1975_2006/, based on November 2008 SEER data submission, posted to the SEER web site, 2009
- Keating NL, Landrum MS, Ayanian JZ, Winer EP, Guadagnoli E. The association of ambulatory care with BC stage at diagnosis among Medicare beneficiaries. *Journal of General Internal Medicine*. 2004; 20:38–44. [PubMed: 15693926]
- Keegan T, John E, Fish K, Alaro-Velcamp T, Clarke C, Gomez S. Breast Cancer Incidence Patterns among California Hispanic Women: Differences by Nativity and Residence in an Enclave. *Cancer Epidemiol Biomarkers Prev*. 2010; 19(5):1208–18. [PubMed: 20447917]
- Kirsner RS, Ma F, Fleming L, Trapido E, Duncan R, Federman DG, Wilkinson JD. The effect of Medicare health care systems on women with breast and cervical cancer. *Obstetrics & Gynecology*. 2005; 105:1381–1388. [PubMed: 15932833]
- Kramer M, Hogue C. Is Segregation Bad For Your Health? *Epidemiologic Reviews*, v. 2009; 31:178–194.
- Lane PW, Nelder JA. Analysis of Covariance and Standardization as Instances of Prediction. *Biometrics*. 1982; 38:613–621. [PubMed: 7171691]
- Laveist TA. The political empowerment and health-status of African-Americans—mapping a new territory. *AJS*. 1992; 97(4):1080–1095.
- Laveist TA. Segregation, poverty, and empowerment: health consequences for African Americans. *Milbank Q*. 1993; 71(1):41–64. [PubMed: 8450822]
- Lee-Feldstein A, Feldstein PJ, Buchmueller T, Katterhagen G. Breast cancer outcomes among older women: HMO, fee-for-service, and delivery system comparisons. *Journal of General Internal Medicine*. 2001; 16:189–199. [PubMed: 11318915]
- MacKinnon J, Duncan R, Huang Y, et al. Detecting an association between socioeconomic status and late stage breast cancer using spatial analysis and area-based measure. *Cancer Epidemiol Biomarker Prev*. 2007; 16:756–762.
- Massey DS, Denton NA. The dimensions of residential segregation. *Social Forces*. 1988; 67:281–315.
- MEDPAC. Dual eligible beneficiaries: An overview. Report to the Congress: New Approaches in Medicare. 2004; Chapter 3
- McCarthy EP, Burns RB, Coughlin SS, Freund KM, Rice J, Marwill SL, Ash A, Shwartz M, Moskowitz MA. Mammography use helps to explain differences in breast cancer stage at diagnosis between older black and white women. *Annals of Internal Medicine*. 1998; 128:729–736. [PubMed: 9556466]
- McElroy JA, Remington PL, Gangnon RE, Hariharan L, Andersen LD. Identifying geographic disparities in the early detection of breast cancer using a geographic information system. *Preventing Chronic Disease*. 2006; 3:A10. [PubMed: 16356363]
- Mobley LR, Kuo TM, Andrews L. How sensitive are multilevel regression findings to defined area of context?: a case study of mammography use in California. *Medical Care Research and Review*. 2008; 65:315–337. [PubMed: 18259047]
- Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Social Science & Medicine*. 2004; 58:1929–1952. [PubMed: 15020009]

- Osypuk T, Acevedo-Garcia D. Beyond individual neighborhoods: A geography of opportunity perspective for understanding racial/ethnic health disparities. *Health & Place*. 2010;10.1016/j.healthplace.2010.07.002
- Randolph WM, Goodwin JS, Mahnken JD, Freeman JL. Regular mammography use is associated with elimination of age-related disparities in size and stage of breast cancer at diagnosis. *Annals of Internal Medicine*. 2002; 137:783–790. [PubMed: 12435214]
- Riley GF, Potosky AL, Klabunde CN, Warren JL, Ballard-Barbash R. Stage at diagnosis and treatment patterns among older women with breast cancer: an HMO and fee-for-service comparison. *Journal of the American Medical Association*. 1999; 281:720–726. [PubMed: 10052442]
- Riley GF, Warren JL, Potosky AL, Klabunde CN, Harlan LC, Osswald MB. Comparison of cancer diagnosis and treatment in Medicare fee-for-service and managed care plans. *Medical Care*. 2008; 46:1108–1115. [PubMed: 18815533]
- Roche LM, Skinner R, Weinstein RB. Use of a geographic information system to identify and characterize areas with high proportions of distant stage breast cancer. *Journal of Public Health Management Practice*. 2002; 8:26–32.
- RTI. Spatial impact factor data, Version 2. RTI International, Research Triangle Park, NC: SEER; 2009.
- SEER fast stats. accessed on January, 10, 2010 <http://seer.cancer.gov/faststats/selections.php?#Output>

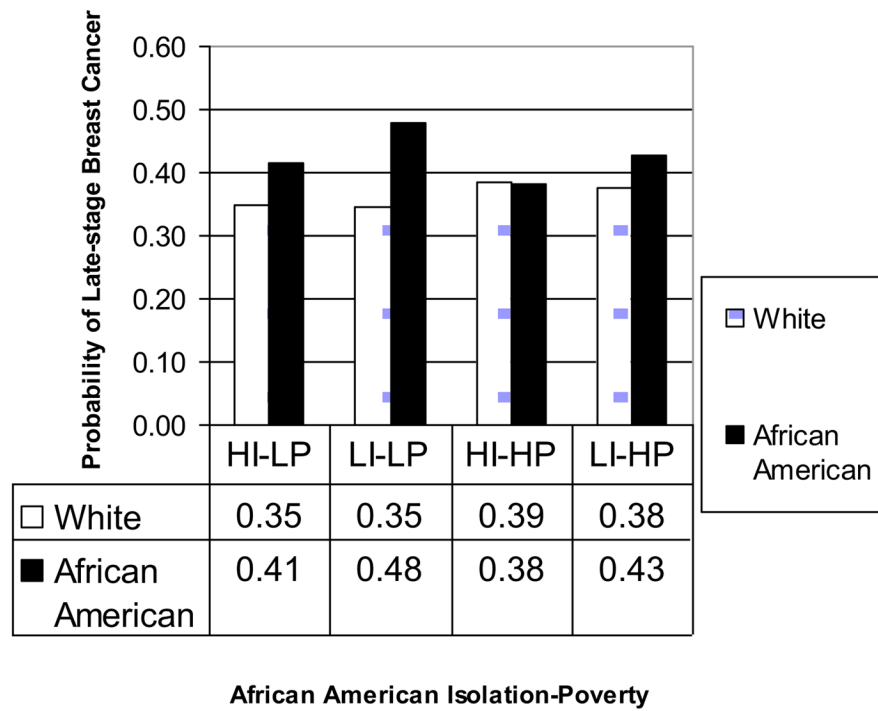


Figure 1. Predicted Marginal Probability of Late-stage Breast Cancer from the African American Model, by African American Isolation Poverty for African American and White Women

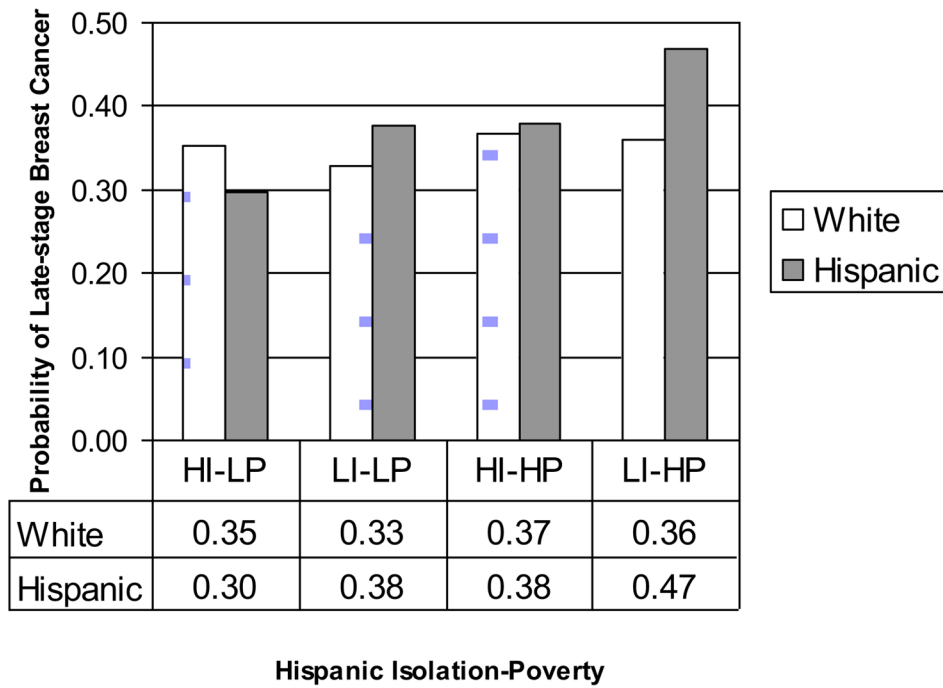


Figure 2. Predicted Marginal Probability of Late-stage Breast Cancer from the Hispanic Model, by Hispanic Isolation–Poverty for Hispanic and White Women

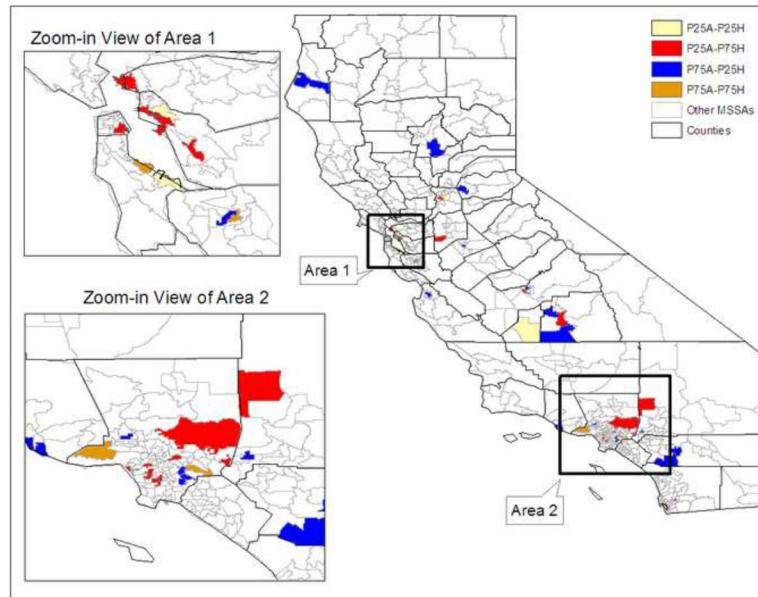


Figure 3.
 Map of Selected MSSA Areas, in Upper and Lower Quartiles of the Distribution of MSSA-level Rates of Late Stage Diagnosis Predicted by the Random Intercepts Multilevel Model
 P25A: below 25th percentile for African Americans
 P75A: above 75th percentile for African Americans
 P25H: below 25th percentile for Hispanics
 P75H: above 75th percentile for Hispanics

Table 1

Sample Characteristics: Personal-level Factors (%)

Variable	White (N=29,693)	African American (N=1,737)	Hispanic (N=1,065)	Asian (N=1,343)
Person-level Variable				
Stage at diagnosis				
Early (localized)	69.1	60.8	62.1	70.7
Late (regional or distant)	30.9	39.2	37.9	29.3
Age groups				
67–69	17.1	20.1	9.2	16.6
70–74	26.7	27.5	30.0	31.6
75–79	25.8	24.6	33.4	28.8
80–84	18.0	16.3	17.5	15.8
85 or older	12.4	11.5	10.0	7.1
Marital status (married)	43.8	27.9	34.6	47.4
State buy-in recipients	10.5	34.3	54.2	50.4
ESRD or disability	0.2	1.4	<1.1 ^a	<1.0 ^a
Any mammography utilization in 2 years prior to diagnosis	27.0	27.7	26.1	30.1

ESRD, End-stage renal disease.

^aN < 11 may not be reported directly or in a derivable way.

Table 2

Characteristics of Area-level Variables (Mean and Standard Deviation Unless Specified Otherwise)

Variable	
MSSA-level data (RTI 2009)	
Isolation index for African American: median (25 th –75 th percentile)	0.030 (0.012 – 0.094)
Isolation index for Hispanics: median (25 th –75 th percentile)	0.274 (0.132 – 0.481)
Isolation Index for Asian: median (25 th –75 th percentile)	0.055 (0.018 – 0.142)
proportion of population aged 65+ living below poverty (25 th –75 th percentile)	0.072 (0.053 – 0.101)
proportion of area is rural	0.232 (0.342)
proportion of workforce that spend more than 60 minutes commuting each way to work	0.104 (0.059)
Mammography facility density (per 1000 females aged 40 and older)	0.051 (0.200)
proportion of population foreign born and immigrated to California 1995–2000	0.175 (0.066)
number of primary care physicians or obstetricians/gynecologists per 1000 population	0.971 (0.652)
County-level Data (Area Resource File)	
Medicare managed care penetration in county, 1999	0.224 (0.190)
Oncologists per 1000 population aged 65+ in county	0.068 (0.066)

Table 3

Coefficient Estimates from Logistic Multilevel Models of Late-stage Breast Cancer: Separate Models by Race and Ethnicity

Person-level Variables	African American and White	Hispanic and White	Asian and White
Age groups			
Age 67–69	ref	ref	ref
Age 70–74	–0.068	–0.075*	–0.081*
Age 75–79	–0.113*	–0.113*	–0.115*
Age 80–84	–0.088*	–0.103*	–0.099*
Age 85 or older	0.062	0.050	0.062
Marital status			
Married	–0.107*	–0.116*	–0.110*
Not married	ref	ref	ref
Race/ethnicity			
African American/Hispanic/Asian	–0.017	–0.260	–0.244*
White	ref	ref	ref
State buy-in recipients			
Yes	0.369*	0.367*	0.392*
No recipients	ref	ref	ref
ESRD or disabled			
Yes	0.227	0.110	0.087
None	ref	ref	ref
Mammography in 2 years prior to diagnosis			
Yes	–0.041	–0.038	–0.039
No	ref	ref	ref
Contextual variables			
Racial isolation-poverty group			
HI-LP	–0.159*	ref	ref
LI-LP	–0.168*	–0.111*	0.013
HI-HP	ref	0.055	0.145*
LI-HP	–0.043	–0.032	0.146*
proportion of area is rural	–0.079	–0.055	–0.103
proportion of workforce commute more than 60 minutes to work	–0.208	–0.110	–0.094
Medicare managed care penetration in 1999 in county	–0.072	–0.047	–0.120
Mammography facility density (per 1000 females aged 40 and older)	0.068	0.079	0.072
proportion of population is foreign born	–0.257	–0.300	–0.217
number of Primary care physicians or obstetricians/gynecologists per 1000 population	–0.029	–0.020	–0.030
Oncologists per 1000 population aged 65+ in county	0.200	0.266	0.188
Interaction Effect of race by income-racial isolation			

Person-level Variables	African American and White	Hispanic and White	Asian and White
Race × HI-LP	0.297 [*]	ref	ref
Race × LI-LP	0.560 [*]	0.475 [*]	0.004
Race × HI-HP	ref	0.309	-0.075
Race × LI-HP	0.226	0.707 [*]	0.145

^{*} Indicates statistical significance at $p < 0.05$.

ESRD, End-stage renal disease; HI-HP, High isolation-high poverty; HI-LP, High isolation-low poverty; LI-LP, Low isolation-low poverty; LI-HP, Low isolation-high poverty.